

9367

From: Bunner, Bridget
Sent: Monday, May 05, 2003 12:36 PM
To: STIC-Biotech/ChemLib
Subject: sequence search

Hi! I'd like to request a sequence search for case 09/894,912 (no pending search is required):

1. the amino acid sequence of SEQ ID NO: 13

Thanks!

Bridget Bunner

Art Unit 1647
CM1-10D12
(703) 305-7148
mailbox 10B19

CRF

Point of Contact

P. Sheppard

Searcher Telephone number: (703) 308-4499

Phone: _____

Location: _____

Date Picked Up: _____

Date Completed: 5/7/03

Searcher Prep/Review: _____

Clerical: _____

Online time: _____

TYPE OF SEARCH:

NA Sequences: _____

AA Sequences: _____

Structures: _____

Bibliographic: _____

Litigation: _____

Full text: _____

Patent Family: _____

Other: _____

VENDOR/COST (where applic.)

STN: _____

DIALOG: _____

Questel/Orbit: _____

DRLink: _____

Lexis/Nexis: _____

Sequence Sys.: _____

WWW/Internet: _____

Other (specify): _____

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein --protein search, using sw model

Run on: May 6, 2003, 14:50:18 ; Search time 33 Seconds
(without alignments)
1704.572 Million cell updates/sec

Title: US-09-894-912A-13

Sequence: 1516
1 MCHLRISWLFILINFMVEYI.....QQKKRKYODKXSVSVTVH 273

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 1505 | 99.3 | 272 | 4 Q9BXV4 | Q9BXV4 homo sapien |
| 2 | 1452 | 95.8 | 292 | 4 Q96K87 | Q96K87 homo sapien |
| 3 | 1153 | 76.1 | 217 | 11 Q9CSB2 | Q9CSB2 mus musculu |
| 4 | 644 | 42.5 | 265 | 11 Q92132 | Q92132 mus musculu |
| 5 | 472 | 31.1 | 224 | 4 Q9UCB2 | Q9UCB2 mus sapien |
| 6 | 192 | 12.7 | 1299 | 5 Q26489 | Q26489 spodioptera |
| 7 | 177 | 11.7 | 1101 | 5 Q964D2 | Q964D2 entameeba h |
| 8 | 176.5 | 11.6 | 1376 | 5 Q8S2S2 | Q8S2S2 drosophila |
| 9 | 176.5 | 11.6 | 1679 | 5 Q24301 | Q24301 drosophila |
| 10 | 176 | 11.3 | 1074 | 5 Q964D1 | Q964D1 entameeba h |
| 11 | 172 | 11.3 | 915 | 11 Q91VK0 | Q91VK0 mus musculu |
| 12 | 167.5 | 11.0 | 296 | 11 Q35171 | Q35171 mus musculu |
| 13 | 167.5 | 11.0 | 932 | 11 Q62030 | Q62030 mus musculu |
| 14 | 162 | 10.7 | 913 | 4 Q96EP4 | Q96EP4 homo sapien |
| 15 | 161 | 10.6 | 440 | 5 Q18003 | Q18003 caenorhabdi |
| 16 | 159 | 10.5 | 503 | 5 Q9U018 | Q9U018 giardia lam |

| | | | | | |
|----|-------|------|------|-----------|--------------------|
| 17 | 158.5 | 10.5 | 379 | 4 Q8WVG4 | Q8WVG4 homo sapien |
| 18 | 157.5 | 10.4 | 374 | 13 Q9W6F8 | Q9W6F8 xenopus lae |
| 19 | 157.5 | 10.4 | 803 | 13 Q42114 | Q42114 brachydanio |
| 20 | 156 | 10.3 | 898 | 5 Q76822 | Q76822 brachydanio |
| 21 | 155 | 10.2 | 808 | 13 Q42113 | Q42113 brachydanio |
| 22 | 153.5 | 10.1 | 942 | 5 Q44762 | Q44762 caenorhabdi |
| 23 | 153.5 | 10.1 | 1362 | 13 Q9PV24 | Q9PV24 xenopus lae |
| 24 | 151 | 10.0 | 288 | 5 Q76510 | Q76510 cryptospori |
| 25 | 150.5 | 9.9 | 827 | 5 Q44765 | Q44765 caenorhabdi |
| 26 | 149 | 9.8 | 1371 | 11 Q9QWV4 | Q9QWV4 rattus sp. |
| 27 | 148 | 9.8 | 548 | 5 Q9GQ45 | Q9GQ45 giardia lam |
| 28 | 147.5 | 9.7 | 802 | 13 Q9W770 | Q9W770 gallus gall |
| 29 | 146.5 | 9.7 | 365 | 11 Q924Y6 | Q924Y6 rattus norv |
| 30 | 146.5 | 9.7 | 807 | 4 Q9HCB6 | Q9HCB6 homo sapien |
| 31 | 145 | 9.6 | 624 | 4 Q94862 | Q94862 homo sapien |
| 32 | 144.5 | 9.5 | 660 | 5 Q21832 | Q21832 cryptospori |
| 33 | 144.5 | 9.5 | 2189 | 5 Q9H105 | Q9H105 elmeria ten |
| 34 | 144 | 9.5 | 807 | 6 Q9GLX9 | Q9GLX9 bos taurus |
| 35 | 143.5 | 9.5 | 220 | 11 Q93KR2 | Q93KR2 mus musculu |
| 36 | 143.5 | 9.5 | 807 | 11 Q8VCC9 | Q8VCC9 mus musculu |
| 37 | 143 | 9.4 | 435 | 11 Q9GQ41 | Q9GQ41 giardia lam |
| 38 | 142 | 9.4 | 378 | 13 Q9W6F9 | Q9W6F9 brachydanio |
| 39 | 142 | 9.4 | 921 | 5 Q969A3 | Q969A3 brachydanio |
| 40 | 141 | 9.3 | 261 | 5 Q95YF9 | Q95YF9 ciona savig |
| 41 | 140 | 9.2 | 259 | 4 Q8WUR5 | Q8WUR5 homo sapien |
| 42 | 138.5 | 9.1 | 1368 | 5 Q21821 | Q21821 caenorhabdi |
| 43 | 138.5 | 9.1 | 1433 | 5 Q9B1H9 | Q9B1H9 anopheles g |
| 44 | 138.5 | 9.1 | 1664 | 5 Q9TVQ2 | Q9TVQ2 caenorhabdi |
| 45 | 137 | 9.0 | 213 | 11 Q99LE4 | Q99LE4 mus musculu |

ALIGNMENTS

RESULT 1

Q9BXV4 PRELIMINARY; PRT; 272 AA.

AC Q9BXV4
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Thrombospondin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Mao Y., Xie Y., Zhou Z., Zhao W., Zhao S., Wang W., Huang Y., Wang S.,
RA Tang R., Chen X., Wu C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RC TISSUE=PLACENTA;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF251057; AAK34947.1;
DR EMBL; BC022367; AAK22367.1;
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp.1; 1.
DR SMART; SM00261; FU. 2.
DR SMART; SM00209; TSP1. 1.
DR PROSITE; PS50092; TSP1. 1.
SQ SEQUENCE 272 AA; 30928 MW; CACACGCB7E781189 CRC64;
Query Match 99.3%; Score 1505; DB 4; Length 272;
Best Local Similarity 100.0%; Pred. No. 9.9e-129;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 HLRISWLFILINFMVEYIGSNASGRGRORRMHPNVSGCGGCAATCSVDYGCSCRPRL 62
DB 2 HLRISWLFILINFMVEYIGSNASGRGRORRMHPNVSGCGGCAATCSVDYGCSCRPRL 61

```

Oy 63 FFALERIGMKOIGVCLSSCPGSGYGTTRYPDINKCTCKADCDCTCNKFCCKCKSGFYHL 122
Db 62 FFALERIGMKOIGVCLSSCPGSGYGTTRYPDINKCTCKADCDCTCNKFCCKCKSGFYHL 121
Oy 123 LGKCLDNCCEGLANNNHMECVSIHCEVSENNPMSPTCKKGTGCFKGTETVRREIIQ 182
Db 122 LGKCLDNCCEGLANNNHMECVSIHCEVSENNPMSPTCKKGTGCFKGTETVRREIIQ 181
Oy 183 HPSAKGNLCPPTNETRKTCTVORRKCCKGGERGKGRERKRRKKNKESKEAI PDSKLSLESS 242
Db 182 HPSAKGNLCPPTNETRKTCTVORRKCCKGGERGKGRERKRRKKNKESKEAI PDSKLSLESS 241
Oy 243 KEIPEORENKOOOKKRRKTVODKOKSVSTVTH 273
Db 242 KEIPEORENKOOOKKRRKTVODKOKSVSTVTH 272

RESULT 2
Oy 096K87 PRELIMINARY; PRT; 292 AA.
AC 096K87;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CNNA FL114440 f15, clone HEMBI000915, weakly similar to
subtilisin-like protease PACE4 precursor (EC 3.4.21.-).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRIO;
RA Itoogi T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
Tanabe T., Nomura Y., Togiyasu S., Kamai F., Hara R., Takeuchi K.,
Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
Makamatsu A., Nakamura Y., Nagahara K., Masuo Y., Ohshima A.;
"NEBO human cDNA sequencing project.";
RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AK027346; BAB55051.1;
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000884; TSP1.
DR Pfam: PRO00090; tsp_1; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS50092; TSP1; 1.
SQ SEQUENCE 292 AA; 33243 MW; 01E2774AC3DA6F8 CRC64;

Query Match 95.8%; Score 1452; DB 4; Length 292;
Best Local Similarity 99.2%; Pred. No. 6, 8e-124;
Matches 262; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

RESULT 3
Oy 09CSB2 PRELIMINARY; PRT; 217 AA.
AC 09CSB2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 2810459H04RIK protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRIO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
Schriml L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Adono H., Balderelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilting L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Koltunski S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
DR EMBL: AK013366; BAB28811.1;
DR MGD; MGI:1920030; 2810459H04RIK.
DR InterPro: IPR002174; Tspn-like.
DR InterPro: IPR000884; TSP1.
DR Pfam: PRO00090; tsp_1; 1.
DR SMART: SM00261; FT2.
DR SMART: SM00269; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
FT NON_TER 217
SQ SEQUENCE 217 AA; 24304 MW; 0DCF93BE9FB3FBF7 CRC64;

Query Match 76.1%; Score 1153; DB 11; Length 217;
Best Local Similarity 93.1%; Pred. No. 6, 6e-97;
Matches 201; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

```

AC 092132;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE Thrombospondin type 1 domain.
 GN R-SPONDIN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kamata T., Katsube K., Michikawa M., Yamada M., Mizusawa H.;
 RT "R-spondin, a novel thrombospondin type 1 domain gene, expressed in
 the dorsal neural tube."
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB016768; BAA75640.1; -
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR000884; TSP1.
 DR Pfam: PF00090; TSP_1.
 DR SMART: SM00261; FU; 2.
 DR SMART: SM00209; TSP1; 1.
 DR PROSITE: PS50092; TSP1; 1.
 SQ SEQUENCE 265 AA; 29331 MW; FEEB8964743F5963 CRC64;
 Query Match 42.5%; Score 644; DB 11; Length 265;
 Best Local Similarity 46.5%; Pred. No. 1,2e-50;
 Matches 119; Conservative 36; Mismatches 77; Indels 24; Gaps 6;
 QY 6 LISWFIINFEVYIGSONASRGRRMRHNPVSGCGGCGATCGDYNGCLSCPRLLFA 65
 DB 11 VLSWTHIA-----VGRGRI-KGRKORRISAEBSQAQAKGCELCSEVNGCLCKSPKLFIL 63
 QY 66 LERIGKKQIGVCLSCSPGYYGTRYPINKCTCKKAD-CDTFNNNPFCTKSKSFYHLG 124
 DB 64 LENDIRRGVCLPCPCPGYDANPNPMKCIKCKIHEEACFSINFTCKQELYLHKG 123
 QY 125 KCLDNPEGLANHTMECVSYHCEVSENPMPSPCTKKGCTGFKRGTETRVREIIGHP 184
 DB 124 RCYPAPEEGSTANSTWCEGSPACCESEWSPGCPCKKRLCGFRKGSERTRRVLHAP 183
 QY 185 SAKNGLCPPTNETRCKTVORKKCKGKRGKQ---REKRRKKPKGSEKALPDSKLE 240
 DB 184 GSDHTTCSPTETKCTYRATPCPEGGKRGKRGQGRNARHAPKNSKE--PRNS-- 239
 QY 241 SSKETPEGRKNOOK 256
 DB 240 -----RRHKGQOQ 247
 RESULT 5
 QYGB2 PRELIMINARY; PRT; 224 AA.
 ID 09UGB2
 AC 09UGB2;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE DJ824F16.3 (Novel protein similar to mouse thrombospondin type 1
 domain protein R-spondin) (Fragment).
 GN DJ824F16.3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Blakey S.;
 RL Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL050325; CAB65783.3; -
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR000884; TSP1.
 DR SMART: SM00261; FU; 2.
 DR SMART: SM00209; TSP1; 1.

DR PROSITE: PS50092; TSP1; 1.
 FT NON_TER 224 224
 SQ SEQUENCE 224 AA; 25042 MW; 97D26AD34CDBFE12 CRC64;
 Query Match 31.1%; Score 472; DB 4; Length 224;
 Best Local Similarity 43.3%; Pred. No. 3,9e-35;
 Matches 87; Conservative 36; Mismatches 70; Indels 8; Gaps 5;
 QY 29 RRRRRMNPVSGCGGCGATCGDYNGCLSCPRLLFALERIGMKQIGVCLSCSPGYYGT 88
 DB 22 RKKQVGTGLGNGCT-GCIISEENGCSCTCOORLFLFRREGIRYQKCLHDCPPGYFGI 80
 QY 89 RYPDINKCTCKADCDCTENKNTCTCKSGFYLHGKCLDNCEGLAEANNHTMECVSYH 148
 DB 81 RQEVNRRKKCGATCSGSCSPFCIRCKRQFYLVKGCILCTCPPTLHONTRREGCG--E 138
 QY 149 CEVSENNPSPCTKKGCTGFKRGTETRVREIIGHPSAKNGLCPPTNETRCKTVQPKKO 208
 DB 139 CELGFGWGSPTCHNGKTCGSAWGLESRVREAGHAEATQVLSERKCPIDR-PC- 196
 QY 209 KGERG--KGRERKKRKNK 226
 DB 197 PDEREPQKKGRKDRPRKDR 217
 RESULT 6
 QYGB2 PRELIMINARY; PRT; 1299 AA.
 ID 026489
 AC 026489;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Endoprotease furin.
 GN FURIN.
 OS Spodoptera frugiperda (Fall armyworm).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Noctuoidea; Noctuidae; Amphipyritinae; Spodoptera.
 NCBI_TaxID=7108;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SP9;
 RT "Cloning and functional characterization of FURIN from Spodoptera
 frugiperda (Sf9) cells."
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z68888; CAA93116.1; -
 DR HSSP: Q99405; IMPT.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR000209; Peptidase_S8.
 DR InterPro: IPR002884; P_domain.
 DR Pfam: PF01483; P_1.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PR00723; SUBTILISTIN.
 DR ProDom: PD000717; P_domain; 1.
 DR SMART: SM00261; FU; 10.
 DR PROSITE: PS00136; SUBTILASE_ASP; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KM Protease.
 SQ SEQUENCE 1299 AA; 142020 MW; 4C3799C7B8C572AB CRC64;
 Query Match 12.7%; Score 192; DB 5; Length 1299;
 Best Local Similarity 27.8%; Pred. No. 5,3e-09;
 Matches 63; Conservative 26; Mismatches 78; Indels 60; Gaps 12;
 QY 37 NVSGCGGCGATCGDYNGCLSCPRLLFALERIGMKQIGVCLSCSPGYYGTRYPINK 95
 DB 795 SVCRPAACATCSEKADGCTSCSHL-----VLHGTGCMACGPFSHET---EDDM 843
 QY 96 CTCKKADCTCF--NKNFTCKKSGFYLHGKCLDNCEGLAEANNHTMECVSYHCEVSE 153
 DB 844 CAKHSCDTCGPGTQCVTCHPSTYALDGRCTVSCPPAYYADKKRKC---MRCPVG- 899

OY 154 MWSPBCTKGGTCGFKGTETRVREIIIOHPSAKNLCPP--TNETKRC-TVORRCKCK 209
 DB 900 ---CSTCT-----SAFCLSCPEKMLKKCKCPVGSDCSA 933
 OY 210 GEGGKGRERKRRKPKN-----GESKE---AIPDSKLSSEKPE 247
 DB 934 GEFAYV---DQKRCRCPACDSCYGENEGHCLTCPPNLLQDYKCPVE 977

RESULT 7

ID Q964D2 PRELIMINARY; PRT; 1101 AA.
 AC Q964D2;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Gal/GalInac lectin Ig11.
 GN IGL1.
 OS Entamoeba histolytica.
 OC Eukaryota; Entamoebidae; Entamoeba.
 OX NCBI_TaxID=5759;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HMI.1MSS;
 RX MEDLINE=21319185; PubMed=11500468;
 RA Cheng X.J., Hughes M.A., Huston C.D., Loftus B., Gilchrist C.A.,
 RA Lockhart L.A., Ghosh S., Miller-Sims V., Mann B.J., Petri W.A. Jr.,
 RA Tachibana H.;
 RT "Intermediate subunit of the Gal/GalInac lectin of Entamoeba
 RT histolytica is a Member of a Gene Family Containing Multiple CXXC
 RT Sequence Motifs";
 RL Infect. Immun. 69:5892-5898 (2001).
 DR EMBL: AF137950; AAK92361.1; -
 DR InterPro: IPR000561; BGF-like
 DR PROSITE: PS01186; BGF_2; UNKNOWN_1.
 KW lectin.
 SQ SEQUENCE 1101 AA; 119512 MW; C8B6F5CBDE656AEC CRC64;

Query Match 11.7%; Score 177; DB 5; Length 1101;
 Best Local Similarity 26.6%; Pred. No. 1e-07;
 Matches 55; Conservative 22; Mismatches 80; Indels 50; Gaps 10;

OY 44 GGCATCSD---YNGCL---SC---KRLFFALF-----RIGKQIGVCLSSCSGYGTR 89
 DB 773 GTGOSCSDSLKPYGCKKTDSCNVDNRGFIYATECSGDSGSPYSNCTTCTKSNY-RK 831
 OY 90 YPDINKCTKCAKADCTCFNKNFCTK-----CKSGFYHLGKCLDNCPEGLEA 136
 DB 832 EGGKNGCAKCDKCATCSDKDTCLTCADPLKVGSKDCKGTYVMSNGEC-----KPC 884
 OY 137 NNHTWCUSIVHCEVSEMNPMSPCTKKG-KTC--GFKGTETRVREIIIOHPSAKNLCPP 193
 DB 885 THHCESSAACTVCSBDTYKYISGNGCNCYVDG-----YDEIKGTCTPC 932
 OY 194 TNETRKTVOVRCKCKGEGKGRERK 220
 DB 933 TSPCTKCVGKKDCEBOETGENSEKKK 959

RESULT 8

ID Q852S2 PRELIMINARY; PRT; 1376 AA.
 AC Q852S2;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE LD30182P.
 GN FUR2.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stepien M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Change M., Chavez C., Dorsett V., Dresnek D., Fairlan D., Fitse E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nuno J., Pacled J., Paragaa V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Ceiniker S.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY070553; AAL48024.1; -
 SQ SEQUENCE 1376 AA; 149716 MW; B6704BA89A3A888B CRC64;

Query Match 11.6%; Score 176.5; DB 5; Length 1376;
 Best Local Similarity 28.0%; Pred. No. 1.4e-07;
 Matches 60; Conservative 24; Mismatches 77; Indels 53; Gaps 12;

OY 3 HURLISWFLILNFMETISQNASRGRMRHNPVSGCGCATCSDY-NGCLSCKPR 61
 DB 734 HLHVLD-LAVCLQCPDGYFENS---RNRTCV-----CEPNCASCDHPREYTSODH 783
 OY 62 LFALERIGMKQIGVCLSCPSGYGTRYPDINKCTKCAKADDTCF--NKVFCCKSGF 119
 DB 784 LVMEHK-----CYSACPDIYET---EDNKAFCSTATCNGPTDODCTCRSSR 832
 OY 120 YLHKGCLDNCPEGLEANNHTWCUSIVHCEVSEMNPMSPCTKKGTCGFKGTETRV-R 178
 DB 833 YAMONKCLISCPDGFYADKKRLECM-----PCDGGCTC-----TSNGVCS 873
 OY 179 EIIHPSAKNLCPPNETRKTVOVRCK-CCKGE 211
 DB 874 ECLDWT-----LNKDKCIVSGSECSSE 899

RESULT 9

ID Q24301 PRELIMINARY; PRT; 1679 AA.
 AC Q24301;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE FUR2 protein.
 GN FUR2 OR DFUR2 OR CG4235 OR CG18734.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Chame M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt R.G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Fiankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,
 RA Fostel A., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Mlishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Xu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*."; *Science* 287:2185-2195(2000).
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ISO-1;
 RX MEDLINE=92361036; PubMed=1512259;
 RA Reebroek A.J., Creemers J.W., Pauli I.G., Kurzik-Dumke U., Rentrop M.,
 RA Gaeft E.A., Leunissen J.A., de Ven W.J.;
 RT "Cloning and functional expression of Dfurin2, a subtilisin-like
 RT proprotein processing enzyme of *Drosophila melanogaster* with multiple
 RT repeats of a cysteine motif."; *J. Biol. Chem.* 267:17208-17215(1992).
 RL J. Biol. Chem. 267:17208-17215(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ISO-1;
 RX MEDLINE=95186060; PubMed=7880443;
 RA Reebroek A.J., Ayoubi T.A., Creemers J.W., Pauli I.G., Ven W.J.;
 RT "The Dfuz2 gene of *Drosophila melanogaster*: genetic organization,
 RT expression during embryogenesis, and pro-protein processing activity
 RT of its translational product Dfuzin2."; *DNA Cell Biol.* 14:223-234(1995).
 RL DNA Cell Biol. 14:223-234(1995).
 DR EMBL: AEO03502; AAF48598.1; -
 DR EMBL: L33831; AAA69860.1; -
 DR HSP: Q89405; IMP1.
 DR MEROPS: S08.049; -
 DR FlyBase: FBgn004596; Fur2.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR002029; Peptidase-S8.
 DR InterPro: IPR002884; P_domain.
 DR Pfam: PF01483; P.1.
 DR Pfam: PF00082; Peptidase-S8; 1.
 DR PRINTS: PR00723; SUBTILISIN.
 DR ProDom: PD000717; P_domain; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00261; FU; 10.
 DR PROSITE: PS00136; SUBTILASE ASP; 1.
 DR PROSITE: PS00137; SUBTILASE HIS; 1.
 DR PROSITE: PS00138; SUBTILASE SER; 1.
 SQ SEQUENCE 1679 AA; 183368 MW; 3F9E749F0B021CF6 CRC64;

Query Match 11.6%; Score 176.5; DB 5; Length 1679;
 Best Local Similarity 28.0%; Pred. No. 1.7e-07;
 Matches 60; Conservative 24; Mismatches 77; Indels 53; Gaps 12;

3 HRLISMLFIILNFMEXYIGSONASRRGRPMHNVNVSOGCGCATGSDY-NGCLSCXKPR 61
 1037 HHLVID-LAVVLQCPDGYFENS---RNRTCVF-----CPNKASQDHPREYCTSCDH 1086

62 LFLAERIGMGIQVCLSCSPGVYGRTPDINKTKRKACDPCF--NKPFTKXSGF 119
 1087 LVMEHK-----CYSACPLOTYET---EDNKAFCSTCATCNGPTDQDITRCSR 1135

120 YHLGKCLDNCPEGLNHNHTMECVSIIVHCEVSENPWSPCTKXGKTGKGTETTRV-R 178
 1136 YAMQKCLISCPDGFVADKKRLCEM-----PQEGCKTC-----TSGVCS 1176

OY 179 EIIHPSAKGNLCPTNTRKTVORKK-CCKGE 211
 1177 ECLQWNT-----LWKDKCLVSGSGCSESE 1202

RESULT 10
 ID 0964D1 PRELIMINARY; PRT; 1074 AA.
 AC 0964D1:
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Gal/GalNAc lectin Ig12 (Fragment).
 GN IGL2.
 OS Entamoeba histolytica.
 OC Eukaryota; Entamoebidae; Entamoeba.
 OX NCBI_TaxID=5759;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HMI:IMSS;
 RX MEDLINE=21391855; PubMed=11500468;
 RA Cheng X.J., Hughes M.A., Huston C.D., Loftus B., Galchrist C.A.,
 RA Lockhart L.A., Ghosh S., Miller-Sims V., Mann B.J., Petri W.A. Jr.,
 RA Tachibana H.;
 RT "Intermediate Subunit of the Gal/GalNAc lectin of *Entamoeba*
 RT histolytica is a Member of the Gal Family Containing Multiple CXXC
 RT Sequence Motifs."; *Infect. Immun.* 69:5892-5898(2001).
 RL EMBL: AF337951; AAK92362.1; -
 DR InterPro: IPR000561; EGF-like.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 FT NON_TER 1 1074
 FT NON_TER 1074 1074
 SQ SEQUENCE 1074 AA; 116782 MW; 1552E2D714EB450F CRC64;

Query Match 11.6%; Score 176; DB 5; Length 1074;
 Best Local Similarity 25.6%; Pred. No. 1.3e-07;
 Matches 53; Conservative 23; Mismatches 81; Indels 50; Gaps 9;

44 GGCATCSD---YNGC-----LSCKPLFPALF---RIGMKOICVCLSCSPGVYGR 89
 757 GTCQSCDLSKYPCKTDTTCNVDSRTGYIATGCSDFSGSPSYNCTGIESNY-PK 815

90 YPDINKTKRKADCDPCFNKNF-----CTKCKSGFYHLGKCLDNCPEGLEA 136
 816 EGEKNGCAKCDKCATCSDKDTCTDTPLKIGSKCDECKTGYSNNEC-----KPC 868

137 NNHTMECVSIIVHCEVSENPWSPCTKXG-KTC--GPKRGTTREVEIIQHPSAKGNLCPP 193
 869 TNHSCSECSAAECTVCESDPYKVISGNCNACVDGF-----YFDEINGTCIPC 916

194 TNERKCTVORKKCKCKGGRKKGRERK 220
 917 TSPCTKCVGVKDCDCEQFTGCNSEKKK 943

RESULT 11
 ID 091VK0 PRELIMINARY; PRT; 915 AA.
 AC 091VK0:
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Unknown (protein for MGC:18501).
 GN PCSK5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BREAST TUMOR;

QY 19 YIGSONASRGRRORRMHPNVSGCGGATGSDYNG--GLSCRPRLFALERIGMKQIGV 76
 DB 700 YFGAARARRRR-----CHKGETCTGRSPACLSGR-RGFY-----HHQENT 742
 QY 77 CLSSCPGYYGTRYPDINKCTCKADCDTCFNK-NFTCKCKSGFYHLGKCLDNCPEGLE 135
 DB 743 CVTLCPAGLVADESORL--CLRCHPSCKCKVDEPEKCTVKEKESFSLARSCIPDEPGTY 800
 QY 136 ANNHTEC-----VSIVHCEVS-----EMNWPSPCTKKGTCGFKGTET 176
 DB 801 FDSLVLVCGCHHTCTCTGVPSPRECHICAKSFHFMQKVPACGE-----GF----- 848
 QY 177 VREIIQHPSAKGNLCPTNETKCTVQKCKQKGER 212
 DB 849 -----YPEMPGL--PKNVCRCECNCLSCGSSR 876

RESULT 14

Q96EP4 * PRELIMINARY; PRT: 913 AA.
 AC Q96EP4 * 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Proteolysin convertase subtilisin/kexin type 5.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NX NCBI_TaxId=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RA Strausberg R.;
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC012064; AAH12064.1; -
 DR MEROPS: S08_076; -
 DR InterPro: IPR000561; EGF-like
 DR InterPro: IPR000209; Peptidase_S8.
 DR InterPro: IPR002884; P_domain.
 DR Pfam: PF01483; P; 1.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR ProDom: PD000717; P_domain; 1.
 DR SMART: SMO0181; EGF; 4.
 DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.
 DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
 DR PROSITE: PS00138; SUBTILASE_SER; UNKNOWN_1.
 SO SEQUENCE 913 AA; 101649 MW; 7A2E63EFC49104BC CRC64;

Query Match 10.7%; Score 162; DB 4; Length 913;

Best Local Similarity 22.3%; Pred. No. 2e-06;

Matches 55; Conservative 33; Mismatches 81; Indels 78; Gaps 12;

QY 21 GSONASGRORRMHPNVSGCGGATGSDYNG--CLSCRPRLFALERIGMKQIGV 78
 DB 677 GHYHAK-KRCKKCAPN-----CESCFGSHDDQMSCKGYFL-----NEETNSCV 721
 QY 79 SSCPSG-YGTRYPDINKCTCKADCDTCFNK-NFTCKCKSGFYHLGKCLDNCPEGLE 137
 DB 722 THCPDGSYDPTK--NLCKKCSKCKTCTEFHNCTEGRDGLSOSRCSVSCEDGRYFN 778
 QY 138 NHTE-----CVS-----IVHCEVSEW-----NPMSPCTK 163
 DB 779 GDCOPCHRFATCAGAGADGCTNCTEGYFMEDRCVQSCSISYFIDSSENGYSCKKC 838
 QY 164 GKTG-----GFKGTE-----TVREIIQHPSAKGNLCPTNETKCTV 202
 DB 839 DISCLTNGPGFKNCTSPSGYLLDLGMQMGAIKDATEESMAEGFCMLVKNNLC-- 896
 QY 203 QKCKCK 209
 DB 897 ORKVLQ 903

RESULT 15

ID 018003 PRELIMINARY; PRT: 440 AA.
 AC 018003;
 DT 01-JAN-1998 (TRENBLrel. 05, Created)
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE R17.3 protein.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 CC Rhabditidae; Pelodierinae; Caenorhabditis.
 NX NCBI_TaxId=6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Barlow K.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99063613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL: Z92809; CAB07269.1; -
 DR InterPro: IPR001212; Somatomedin_B.
 DR InterPro: IPR000884; TSP1.
 DR Pfam: PF00090; TSP1; 1.
 DR Pfam: PF00209; TSP1; 1.
 DR SMART: SMO0209; TSP1; 1.
 DR PROSITE: PS00524; SOMATOMEDIN_B; 1.
 DR PROSITE: PS50092; TSP1; 1.
 SO SEQUENCE 440 AA; 50400 MW; 42820B834F263EE6 CRC64;

Query Match 10.6%; Score 161; DB 5; Length 440;

Best Local Similarity 23.7%; Pred. No. 1.2e-06;

Matches 53; Conservative 30; Mismatches 87; Indels 54; Gaps 8;

QY 59 KRLFLFALERIGMKQIGVCLSSCPGYYGTRYPDINKCTCKADCDTCFNK-NFTCKCKSG 118
 DB 172 KPRHLLIRYSILSKFMPLKVTSTPLYLEENRVQPMANLYLSSISSEYCDHC----- 225
 QY 119 FYHLGKCLDNCPEGLEANNHTECVSIVHCEVSEPMNPMSPCTKKGTCGFKGTETRYR 178
 DB 226 --VTIGDC-----CSYTRVCP--RDCVLTDMDSWTOCTADNGTGI--GTOKLR 271
 QY 179 EIIQHPSAKGNLCPTNETKCTVQKCKQKGERGKGRERKRRKPKNGESKEAIPDSKS 238
 DB 272 HVIOHARGAACEPLKEMRTCFVE--CR-----PKSALDD-- 306
 QY 239 LESSKEIPEORENKQOKKRX-----VODKOKSVSVYH 273
 DB 307 ITTVALIDYRHNKTRKXIRNNIYWDLPNVAEKKATITYCVH 350

Search completed: May 6, 2003, 14:52:39

Job time : 37 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:51:43 ; Search time 22 seconds
(without alignments)

1070.739 Million cell updates/sec

Title: US-09-894-912A-13

Perfect score: 1516
Sequence: 1 MGLRLISWLFITLNFMEYI.....QOKRRYDQKXSVSTVH 273

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328255 seqs, 86286685 residues

Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------------|
| 1 | 1516 | 100.0 | 273 | 9 | US-09-894-912A-13 |
| 2 | 1505 | 99.3 | 272 | 9 | US-10-125-852-23 |
| 3 | 1505 | 99.3 | 272 | 9 | US-09-894-912A-10 |
| 4 | 1505 | 99.3 | 272 | 9 | US-09-894-912A-34 |
| 5 | 1472 | 97.1 | 265 | 9 | US-09-894-912A-26 |
| 6 | 1472 | 97.1 | 265 | 9 | US-10-125-852-25 |
| 7 | 1472 | 97.1 | 265 | 9 | US-09-894-912A-48 |
| 8 | 1472 | 97.1 | 265 | 9 | US-09-745-763-166 |
| 9 | 1400 | 92.3 | 251 | 9 | US-09-894-912A-16 |
| 10 | 1310.5 | 86.4 | 229 | 9 | US-09-894-912A-32 |
| 11 | 1275 | 84.1 | 225 | 9 | US-10-185-770-4 |
| 12 | 903 | 59.6 | 160 | 9 | US-09-894-912A-14 |
| 13 | 656 | 43.3 | 263 | 9 | US-10-125-852-18 |
| 14 | 648.5 | 42.8 | 243 | 9 | US-10-125-852-21 |
| 15 | 644 | 42.5 | 243 | 9 | US-10-125-852-24 |
| 16 | 638 | 42.1 | 229 | 9 | US-09-894-912A-25 |
| 17 | 584.5 | 38.6 | 243 | 9 | US-10-185-770-2 |
| 18 | 577.5 | 38.1 | 243 | 9 | US-10-125-852-13 |
| 19 | 573 | 37.8 | 250 | 9 | US-10-125-852-3 |

| | | | | | | |
|----|-------|------|-----|----|-------------------|-------------------|
| 20 | 545 | 35.9 | 222 | 9 | US-10-125-852-15 | Sequence 15, Appl |
| 21 | 540.5 | 35.7 | 229 | 9 | US-10-125-852-6 | Sequence 6, Appl |
| 22 | 276 | 18.2 | 46 | 9 | US-09-894-912A-18 | Sequence 18, Appl |
| 23 | 234.5 | 15.5 | 131 | 9 | US-10-125-852-9 | Sequence 9, Appl |
| 24 | 221 | 14.6 | 42 | 9 | US-09-894-912A-22 | Sequence 22, Appl |
| 25 | 213 | 14.1 | 37 | 9 | US-09-894-912A-20 | Sequence 20, Appl |
| 26 | 202 | 13.3 | 110 | 9 | US-10-125-852-11 | Sequence 11, Appl |
| 27 | 177.5 | 11.7 | 43 | 9 | US-10-125-852-7 | Sequence 7, Appl |
| 28 | 173.5 | 11.4 | 969 | 9 | US-09-961-403-6 | Sequence 6, Appl |
| 29 | 160.5 | 10.6 | 479 | 10 | US-09-764-898-221 | Sequence 221, App |
| 30 | 158.5 | 10.5 | 379 | 9 | US-09-905-291A-4 | Sequence 4, Appl |
| 31 | 158.5 | 10.5 | 379 | 9 | US-10-066-500-22 | Sequence 22, Appl |
| 32 | 158.5 | 10.5 | 379 | 9 | US-09-902-853-4 | Sequence 4, Appl |
| 33 | 158.5 | 10.5 | 379 | 9 | US-09-907-824-4 | Sequence 4, Appl |
| 34 | 158.5 | 10.5 | 379 | 9 | US-09-907-841-4 | Sequence 4, Appl |
| 35 | 158.5 | 10.5 | 379 | 9 | US-09-904-011-4 | Sequence 4, Appl |
| 36 | 158.5 | 10.5 | 379 | 9 | US-10-028-072-308 | Sequence 308, App |
| 37 | 158.5 | 10.5 | 379 | 9 | US-09-906-742-4 | Sequence 4, Appl |
| 38 | 158.5 | 10.5 | 379 | 9 | US-10-121-048-308 | Sequence 308, App |
| 39 | 158.5 | 10.5 | 379 | 9 | US-10-123-904-308 | Sequence 308, App |
| 40 | 158.5 | 10.5 | 379 | 9 | US-10-140-470-308 | Sequence 308, App |
| 41 | 158.5 | 10.5 | 379 | 9 | US-09-906-838-4 | Sequence 4, Appl |
| 42 | 158.5 | 10.5 | 379 | 9 | US-09-907-613-4 | Sequence 4, Appl |
| 43 | 158.5 | 10.5 | 379 | 9 | US-09-907-942-4 | Sequence 4, Appl |
| 44 | 158.5 | 10.5 | 379 | 9 | US-10-175-746-308 | Sequence 308, App |
| 45 | 158.5 | 10.5 | 379 | 9 | US-10-176-918-308 | Sequence 308, App |

ALIGNMENTS

```
RESULT 1
US-09-894-912A-13
; Sequence 13, Application US/09894912A
; Publication No. US20030044792A1
; GENERAL INFORMATION:
; APPLICANT: Tang et al.
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH
; FILE REFERENCE: 28110/37260A
; CURRENT APPLICATION NUMBER: US/09/894,912A
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: To be assigned
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/266,614
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/215,733
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 09/757,562
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 09/543,774
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 13
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-894-912A-13
Query Match 100.0%; Score 1516; DB 9; Length 273;
Best Local Similarity 100.0%; Pred. No. 7.6e-101;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLRLISWLFITLNFMEYISQNASGRORRMPNVSQCGGCATCSQYNCLSCKP 60
DB 1 MGLRLISWLFITLNFMEYISQNASGRORRMPNVSQCGGCATCSQYNCLSCKP 60
QY 1 RLPFALEIRIGMQIVCSCSPSGYGYTRPDINKTKCKADCTCFNKNFCTCKSGGFY 120
DB 61 RLPFALEIRIGMQIVCSCSPSGYGYTRPDINKTKCKADCTCFNKNFCTCKSGGFY 120
QY 121 LHLGKCLDNCPEGLANNTMECVSIHCEVSEMPNWSPTCKKXKTCFCKGTETRVREI 180
```

```
Db 121 LHGKLDNCEPGLANNTMECVSIHCEVSENNPMSPTCKGTCGFRGTEIVREI 180
; PRIOR APPLICATION NUMBER: 60/266,614
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/215,733
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 09/757,562
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 09/543,774
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-894-912a-10

RESULT 2
US-10-125-852-23
; Sequence 23, Application US/10125852
; Publication No. US20030032034A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH FACTOR-1
; FILE REFERENCE: HVS-43A
; CURRENT APPLICATION NUMBER: US/10/125,852
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: US 60/316,368
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-125-852-23

Query Match 99.3%; Score 1505; DB 9; Length 272;
Best Local Similarity 100.0%; Pred. No. 4.6e-100;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 HLRILSWLFIILNFMETIGSONASRGRRORRHPNVSOCCGGCATCSDYNGCLSCKPRL 62
Db 2 HLRILSWLFIILNFMETIGSONASRGRRORRHPNVSOCCGGCATCSDYNGCLSCKPRL 61
Qy 63 FFLALRIGMKQIGVCLSSCPGSGYIGTRYPDINKCTKACDCTCENKPFCTKCKSGFYIH 122
Db 62 FFLALRIGMKQIGVCLSSCPGSGYIGTRYPDINKCTKACDCTCENKPFCTKCKSGFYIH 121
Qy 123 LGKLDNCEPGLANNTMECVSIHCEVSENNPMSPTCKGTCGFRGTEIVREI 182
Db 122 LGKLDNCEPGLANNTMECVSIHCEVSENNPMSPTCKGTCGFRGTEIVREI 181
Qy 183 HPSAKGNLCPTNETRKTCTVQRKKCKGGERKKRKRKPKNGESKEALPDSKLESS 242
Db 182 HPSAKGNLCPTNETRKTCTVQRKKCKGGERKKRKRKPKNGESKEALPDSKLESS 241
Qy 243 KEIPEORENKQOKKRVODKOKSVSVTVH 273
Db 242 KEIPEORENKQOKKRVODKOKSVSVTVH 272

RESULT 3
US-09-894-912a-10
; Sequence 10, Application US/09894912A
; Publication No. US20030044792A1
; GENERAL INFORMATION:
; APPLICANT: Tang et al.
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH
; FILE REFERENCE: 28110/37260A
; CURRENT APPLICATION NUMBER: US/09/894,912A
; CURRENT FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: To be assigned
; PRIOR FILING DATE: 2001-04-05
```

```
; PRIOR APPLICATION NUMBER: 60/266,614
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/215,733
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 09/757,562
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 09/543,774
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-894-912a-10

Query Match 99.3%; Score 1505; DB 9; Length 272;
Best Local Similarity 100.0%; Pred. No. 4.6e-100;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 HLRILSWLFIILNFMETIGSONASRGRRORRHPNVSOCCGGCATCSDYNGCLSCKPRL 62
Db 2 HLRILSWLFIILNFMETIGSONASRGRRORRHPNVSOCCGGCATCSDYNGCLSCKPRL 61
Qy 63 FFLALRIGMKQIGVCLSSCPGSGYIGTRYPDINKCTKACDCTCENKPFCTKCKSGFYIH 122
Db 62 FFLALRIGMKQIGVCLSSCPGSGYIGTRYPDINKCTKACDCTCENKPFCTKCKSGFYIH 121
Qy 123 LGKLDNCEPGLANNTMECVSIHCEVSENNPMSPTCKGTCGFRGTEIVREI 182
Db 122 LGKLDNCEPGLANNTMECVSIHCEVSENNPMSPTCKGTCGFRGTEIVREI 181
Qy 183 HPSAKGNLCPTNETRKTCTVQRKKCKGGERKKRKRKPKNGESKEALPDSKLESS 242
Db 182 HPSAKGNLCPTNETRKTCTVQRKKCKGGERKKRKRKPKNGESKEALPDSKLESS 241
Qy 243 KEIPEORENKQOKKRVODKOKSVSVTVH 273
Db 242 KEIPEORENKQOKKRVODKOKSVSVTVH 272

RESULT 4
US-09-894-912a-34
; Sequence 34, Application US/09894912A
; Publication No. US20030044792A1
; GENERAL INFORMATION:
; APPLICANT: Tang et al.
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH
; FILE REFERENCE: 28110/37260A
; CURRENT APPLICATION NUMBER: US/09/894,912A
; CURRENT FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: To be assigned
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/266,614
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/215,733
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 09/757,562
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 09/543,774
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-894-912a-34

Query Match 99.3%; Score 1505; DB 9; Length 272;
Best Local Similarity 100.0%; Pred. No. 4.6e-100;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

| | | | |
|----|-----|---|-----|
| Qy | 3 | HLHLSIMLEITLINTMEYIGSONSRRORRRNHPVIOGCGGACATSPYNGCSCRPRL | 62 |
| Db | 2 | HLHLSIMLEITLINTMEYIGSONSRRORRRNHPVIOGCGGACATSPYNGCSCRPRL | 61 |
| Qy | 63 | FPLALBIEIGKIOGVCLSSCPSGYGYGRYPDINCKTKKADCDPTCFPNKJCFKCKSGFYLH | 122 |
| Db | 62 | FPLALBIEIGKIOGVCLSSCPSGYGYGRYPDINCKTKKACDPTCFPNKJCFKCKSGFYLH | 121 |
| Qy | 123 | LKGCIDNCPGLEJLANNHTMECVSIVHCEVSENNPMASCTCKKGTGCFKRGTEVREITQ | 182 |
| Db | 122 | LKGCIDNCPGLEJLANNHTMECVSIVHCEVSENNPMASCTCKKGTGCFKRGTEVREITQ | 181 |
| Qy | 183 | HSPAKNCLCPPINETRCKTYQKCKCKCGERKCKGERKPKKPKGSKSEALIPDSKLESS | 242 |
| Db | 182 | HSPAKNCLCPPINETRCKTYQKCKCKCGERKCKGERKPKKPKGSKSEALIPDSKLESS | 241 |
| Qy | 243 | KELPEORENKDOOKKRRVODKOKSVSVSTVH | 273 |
| Db | 242 | KELPEORENKDOOKKRRVODKOKSVSVSTVH | 272 |

```

RESULT 5
US-09-894-912A-26
; Sequence 26, Application US/09894912A
; Publication No. US20030044792A1
; GENERAL INFORMATION:
; APPLICANT: Tang et al.
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH
; FILE REFERENCE: 28110/37260A
; CURRENT APPLICATION NUMBER: US/09/894,912A
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: To be assigned
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/266,614
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/215,733
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 09/757,562
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 09/543,774
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 265
; TYPE: prt
; ORGANISM: Homo sapiens
; US-09-894-912A-26

```

```

Query Match# 97.1%; Score 1472; DB 9; Length 265;
Best Local Similarity 100.0%; Pred. No. 9, 9e-98;
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 HLRISMLPTILNMEYIGSNASRGRORRMHNVNCGCGGCATGCDNYGCLSCAPRL 62
Db 2 HLRHLSMLPTILNMEYIGSNASRGRORRMHNVNCGCGGCATGCDNYGCLSCAPRL 61

Oy 63 FFALEIRIGMKOIGVCLSSCPSGYGYTRYPDINKCTKCADCTCNKNFCTKCSGFYLH 122
Db 62 FFALEIRIGMKOIGVCLSSCPSGYGYTRYPDINKCTKCADCTCNKNFCTKCSGFYLH 121

Oy 123 LKGLCLDNCPEGLLEANNHTMECVSIYHCEVSEMNWMSFCTKGGTGGFRGRTETRVREII 182
Db 122 LKGLCLDNCPEGLLEANNHTMECVSIYHCEVSEMNWMSFCTKGGTGGFRGRTETRVREII 181

Oy 183 HSAAGKNLCPTNNTBTRCTYQVRKCKCKEPRKKRPHKESKEALPDSKLSLSS 242
Db 182 HSAAGKNLCPTNNTBTRCTYQVRKCKCKEPRKKRPHKESKEALPDSKLSLSS 241

Oy 243 KEIPEORENNKQOQKKRRVDDKQKS 266
Db 243 KEIPEORENNKQOQKKRRVDDKQKS 266

```

Db 242 KIPEORENKQOKKRVODKXS 265

RESULT 6
US-10-125-852-25
Sequence 25, Application US/10125852
Publication No. US20030032034A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH FACTOR
TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES
FILE REFERENCE: HVS-43A
CURRENT APPLICATION NUMBER: US/10/125,852
CURRENT FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: US 60/316,368
PRIOR FILING DATE: 2001-08-30
PRIOR APPLICATION NUMBER: US 09/799,451
PRIOR FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.1
SEQ ID NO 25
LENGTH: 292
TYPE: prt
ORGANISM: Homo sapiens
US-10-125-852-25

| | | | | |
|-----------------------|----------------|---|----------|------------|
| Query Match | 97.1% | Score 1472 | DB 9 | Length 292 |
| Best Local Similarity | 100.0% | Pred. No. 1.e-97 | | |
| Matches 264 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |
| Qy | 3 | HLRLISLFLITLNMETIGSONASRGRORRMHNNVSOGCGGACATCDNYNGLSCS | PRL | 62 |
| Db | 2 | HLRLISLFLITLNMETIGSONASRGRORRMHNNVSOGCGGACATCDNYNGLSCS | PRL | 61 |
| Qy | 63 | FFALERIMKOIGVCLSCPSGYYGTRPDIINKCTKAKADCDTCENKPFCTKCSGYLH | | 122 |
| Db | 62 | FFALRIRIMKOIGVCLSCPSGYGTRPDIINKCTKAKADCDTCENKPFCTKCSGYLH | | 121 |
| Qy | 123 | LKCLICDNCPEBLEANNHTMECSYHACVSSNNPMSCTYKKGKTGFGFRGETVRRIIQ | | 182 |
| Db | 122 | LKCLICDNCPEBLEANNHTMECSYHACVSSNNPMSCTYKKGKTGFGFRGETVRRIIQ | | 181 |
| Qy | 183 | HPSAKGNLCPTNETETKCTVORRKKQCKEERKKQREKAKKPNKGSKEALPDSKSLSS | | 242 |
| Db | 182 | HPSAKGNLCPTNETETKCTVORRKKQCKEERKKQREKAKKPNKGSKEALPDSKSLSS | | 241 |
| Qy | 243 | KEIPBQRENKKQCKKRRVDDKXS | | 266 |
| Db | 242 | KEIPBQRENKKQCKKRRVDDKXS | | 265 |

RESULT 7
 US-09-894-912A-48
 Sequence 48, Application US/09894912A
 Publication No. US20030044792A1
 GENERAL INFORMATION:
 APPLICANT: Tang et al.
 TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH
 TITLE OF INVENTION: FACTOR-LIKE POLYPEPTIDES AND POLYNUCLEOTIDES
 FILE REFERENCE: 28110/37260A
 CURRENT APPLICATION NUMBER: US/09/894,912A
 CURRENT FILING DATE: 2002-05-10
 PRIOR APPLICATION NUMBER: To be assigned
 PRIOR FILING DATE: 2001-04-05
 PRIOR APPLICATION NUMBER: 60/266,614
 PRIOR FILING DATE: 2001-02-05
 PRIOR APPLICATION NUMBER: 60/215,713
 PRIOR FILING DATE: 2000-06-28
 PRIOR APPLICATION NUMBER: 09/757,562
 PRIOR FILING DATE: 2001-01-09
 PRIOR APPLICATION NUMBER: 09/543,774
 PRIOR FILING DATE: 2000-04-05
 NUMBER OF SEQ ID NOS: 48

SOFTWARE: Patentin version 3.0
SEQ ID NO 48
LENGTH: 292
TYPE: PRT
ORGANISM: Homo sapiens
US-09-894-912a-48

Query Match 97.1%; Score 1472; DB 9; Length 292;
Best Local Similarity 100.0%; Pred. No. 1.1e-97;
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HLRISWLFILINFEYIGSONASRGRRMRHNPVNSOGCGCATCSDYNGCLSKPRL 62
DB 2 HLRISWLFILINFEYIGSONASRGRRMRHNPVNSOGCGCATCSDYNGCLSKPRL 61
QY 63 FFALEIRGMQIGVCLSSCPGSGYGTIRYPINCKTCKADCDTCNNKPFCTKCKSGFYHL 122
DB 62 FFALEIRGMQIGVCLSSCPGSGYGTIRYPINCKTCKADCDTCNNKPFCTKCKSGFYHL 121
QY 123 LGKCLDNCPEGLEANNHTMECVSIHCEVSENNPWSPTCKGKTGCFKRGTEIRVREIIQ 182
DB 122 LGKCLDNCPEGLEANNHTMECVSIHCEVSENNPWSPTCKGKTGCFKRGTEIRVREIIQ 181
QY 183 HPSAKGNLCPTNETRKTCTVORRCKQGERGKGRERKRRKPKNGESKEALPDSKLESS 242
DB 182 HPSAKGNLCPTNETRKTCTVORRCKQGERGKGRERKRRKPKNGESKEALPDSKLESS 241
QY 243 KEIPEORENKQOQKKRKYODKOKS 266
DB 242 KEIPEORENKQOQKKRKYODKOKS 265

RESULT 8
US-09-745-763-166
Sequence 166, Application US/09745763
Patent No. US20020065394A1
GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth
McCoy, John M.
Lavallie, Edward R.
Collins-Racie, Lisa A.
Evans, Cheryl
Merberg, David
Trecay, Maurice
Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 219
CORRESPONDENCE ADDRESS:
ADDRESS: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/745,763
FILING DATE: 18-Jun-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,333
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 166:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids

TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 166:
US-09-745-763-166

Query Match 97.1%; Score 1472; DB 10; Length 292;
Best Local Similarity 100.0%; Pred. No. 1.1e-97;
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HLRISWLFILINFEYIGSONASRGRRMRHNPVNSOGCGCATCSDYNGCLSKPRL 62
DB 2 HLRISWLFILINFEYIGSONASRGRRMRHNPVNSOGCGCATCSDYNGCLSKPRL 61
QY 63 FFALEIRGMQIGVCLSSCPGSGYGTIRYPINCKTCKADCDTCNNKPFCTKCKSGFYHL 122
DB 62 FFALEIRGMQIGVCLSSCPGSGYGTIRYPINCKTCKADCDTCNNKPFCTKCKSGFYHL 121
QY 123 LGKCLDNCPEGLEANNHTMECVSIHCEVSENNPWSPTCKGKTGCFKRGTEIRVREIIQ 182
DB 122 LGKCLDNCPEGLEANNHTMECVSIHCEVSENNPWSPTCKGKTGCFKRGTEIRVREIIQ 181
QY 183 HPSAKGNLCPTNETRKTCTVORRCKQGERGKGRERKRRKPKNGESKEALPDSKLESS 242
DB 182 HPSAKGNLCPTNETRKTCTVORRCKQGERGKGRERKRRKPKNGESKEALPDSKLESS 241
QY 243 KEIPEORENKQOQKKRKYODKOKS 266
DB 242 KEIPEORENKQOQKKRKYODKOKS 265

RESULT 9
US-09-894-912a-16
Sequence 16, Application US/09894912A
Publication No. US20030044792A1
GENERAL INFORMATION:

APPLICANT: Tang et al.
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH
FILE REFERENCE: 28110/37260A
CURRENT APPLICATION NUMBER: US/09/894,912A
CURRENT FILING DATE: 2002-05-10
PRIOR APPLICATION NUMBER: To be assigned
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/266,614
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/215,733
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 09/757,562
PRIOR FILING DATE: 2001-01-09
PRIOR APPLICATION NUMBER: 09/543,774
PRIOR FILING DATE: 2000-04-05
NUMBER OF SEQ ID NOS: 48
SOFTWARE: Patentin version 3.0
SEQ ID NO 16
LENGTH: 251
TYPE: PRT
ORGANISM: Homo sapiens
US-09-894-912a-16

Query Match 92.3%; Score 1400; DB 9; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.3e-92;
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 QNARSGRRMRHNPVNSOGCGCATCSDYNGCLSKPRLFFALEIRGMQIGVCLSSCP 82
DB 1 QNARSGRRMRHNPVNSOGCGCATCSDYNGCLSKPRLFFALEIRGMQIGVCLSSCP 60
QY 83 SGYGTIRYPINCKTCKADCDTCNNKPFCTKCKSGFYHLGKCLDNCPEGLEANNHTME 142
DB 61 SGYGTIRYPINCKTCKADCDTCNNKPFCTKCKSGFYHLGKCLDNCPEGLEANNHTME 120

QY 156 PMSPTKAGTCGKAGTETREVEIIOHPSAKGNLCPTNETRKTCTVORRCKOGERCK 215
 DB 61 PMSPTKAGTCGKAGTETREVEIIOHPSAKGNLCPTNETRKTCTVORRCKOGERCK 120
 QY 216 GREKRRKPKKESKALIPDSKLSLESSKEIPEORENKOQ 255
 DB 121 GREKRRKPKKESKALIPDSKLSLESSKEIPEORENKOQ 160

RESULT 13
 US-10-125-852-18
 ; Sequence 18, Application US/10125852
 ; Publication No. US20030032034A1
 ; GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
 TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH FACTOR-1
 FILE REFERENCE: HYS-43A
 CURRENT APPLICATION NUMBER: US/10/125,852
 CURRENT FILING DATE: 2002-08-20
 PRIOR APPLICATION NUMBER: US 60/316,368
 PRIOR FILING DATE: 2001-08-30
 PRIOR APPLICATION NUMBER: US 09/799,451
 NUMBER OF SEQ ID NOS: 25
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 18
 LENGTH: 263
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-125-852-18

Query Match 43.3%; Score 656; DB 9; Length 263;
 Best Local Similarity 45.3%; Pred. No. 1,2e-39;
 Matches 117; Conservative 43; Mismatches 76; Indels 22; Gaps 4;

QY 6 LISMLFIILNMEYIGSONASGRORRHPNVSOGCGCATGSDYNGCLSCPRLPFA 65
 DB 11 VLSTWHTA-----SRGIRKGRORRISAEQSQAACAGCELSGELCKSPKLFIL 63
 QY 66 LERIKMOIGVCLSSCPGSGYGTVPDINKCTCKAD-CDTGNKPFCTCKSGFYHLG 124
 DB 64 LERNDIRQVGVCLPSPGYPFARNPDNKKICIKIEHCACFSHNFCTCKCKEGLYLHKG 123
 QY 125 KCLDNCEGLEANNHTMECVSIHCEVSENNPMSPTCKGKTGFRGTETREVEIIOHP 184
 DB 124 RCPACPEGSSAANGTMECSSPAOCSEMSWPMGPKSKQOLCGFRGSEERTRVLHAP 183
 QY 185 SAKGNLCPTNETRKTCTVORRCKOGERCKGRERKPKKPNKG-----ESKALIPDSKLS 240
 DB 184 VGDHAACTDKETRTCTVRRVPCPEGOKRRKKGGRRENANRLARKESKEAGAGR--- 240
 QY 241 SSKETPEORENKOQKR 258
 DB 241 -----RRKGQOQOQOQ 251

RESULT 14
 US-10-125-852-21
 ; Sequence 21, Application US/10125852
 ; Publication No. US20030032034A1
 ; GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
 TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH FACTOR-1
 FILE REFERENCE: HYS-43A
 CURRENT APPLICATION NUMBER: US/10/125,852
 CURRENT FILING DATE: 2002-08-20
 PRIOR APPLICATION NUMBER: US 60/316,368
 PRIOR FILING DATE: 2001-08-30
 PRIOR APPLICATION NUMBER: US 09/799,451
 PRIOR FILING DATE: 2001-03-05

NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 21
 ; LENGTH: 243
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-125-852-21

Query Match 42.8%; Score 648.5; DB 9; Length 243;
 Best Local Similarity 48.1%; Pred. No. 3.7e-39;
 Matches 114; Conservative 39; Mismatches 69; Indels 15; Gaps 3;

QY 27 RGRORRHPNVSOGCGCATGSDYNGCLSCPRLPFALEIRGMIOIGVCLSSCPGTY 86
 DB 5 KGRORRISAEQSQAACAGCELSGELCKSPKLFILERNDIRQVGVCLPSPGTY 64
 QY 87 GTRYPDINKCTCKAD-CDTGNKPFCTCKSGFYHLGKCLDNCEGLEANNHTMECVS 145
 DB 65 DARNPDNKKICIKIEHCACFSHNFCTCKCKEGLYLHKGRCYPACPEGSSAANGTMECSS 124
 QY 146 IHCEVSENNPMSPTCKGKTGFRGTETREVEIIOHPSAKGNLCPTNETRKTCTVORR 205
 DB 125 PQCSEMSWPMGPKSKQOLCGFRGSEERTRVLHAPVGDHAACTDKETRTCTVRRV 184
 QY 206 KOGERCKGRERKPKKPNKG-----ESKALIPDSKLSLESSKEIPEORENKOQKR 258
 DB 185 PCPEGOKRRKKGGRRENANRLARKESKEAGAGR-----RRKGQOQOQOQ 231

RESULT 15
 US-10-125-852-24
 ; Sequence 24, Application US/10125852
 ; Publication No. US20030032034A1
 ; GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
 TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH FACTOR
 FILE REFERENCE: HYS-43A
 CURRENT APPLICATION NUMBER: US/10/125,852
 CURRENT FILING DATE: 2002-08-20
 PRIOR APPLICATION NUMBER: US 60/316,368
 PRIOR FILING DATE: 2001-08-30
 PRIOR APPLICATION NUMBER: US 09/799,451
 PRIOR FILING DATE: 2001-03-05
 NUMBER OF SEQ ID NOS: 25
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 24
 LENGTH: 265
 TYPE: PRT
 ORGANISM: Mus musculus
 US-10-125-852-24

Query Match 42.5%; Score 644; DB 9; Length 265;
 Best Local Similarity 46.5%; Pred. No. 8.4e-39;
 Matches 119; Conservative 36; Mismatches 77; Indels 24; Gaps 6;

QY 6 LISMLFIILNMEYIGSONASGRORRHPNVSOGCGCATGSDYNGCLSCPRLPFA 65
 DB 11 VLSTWHTA-----VGRGRI-KGRORRISAEQSQAACAGCELSGELCKSPKLFIL 63
 QY 66 LERIKMOIGVCLSSCPGSGYGTVPDINKCTCKAD-CDTGNKPFCTCKSGFYHLG 124
 DB 64 LERNDIRQVGVCLPSPGYPFARNPDNKKICIKIEHCACFSHNFCTCKCKEGLYLHKG 123
 QY 125 KCLDNCEGLEANNHTMECVSIHCEVSENNPMSPTCKGKTGFRGTETREVEIIOHP 184
 DB 124 RCPACPEGSSAANGTMECSSPAOCSEMSWPMGPKSKKOLCGFRGSEERTRVLHAP 183
 QY 185 SAKGNLCPTNETRKTCTVORRCKOGERCKGK-----RERKPKKESKALIPDSKLS 240
 DB 184 GGDHTTSDDTKETRTCTVRRVPCPEGOKRRKKGGRRENANRLARKESKEAGAGR--PRSNS-- 239
 QY 241 SSKETPEORENKOQKR 256

Db 240 -----RRHKG000 247

Search completed: May 6, 2003, 14:53:51
Job time : 23 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:50:58 ; Search time 15 seconds
(without alignments)

535,497 Million cell updates/sec

Title: US-09-894-912A-13

Perfect score: 1516
Sequence: 1 MGLRLISWLFILNFMET.....QQRKRVQKXSVSTVH 273

Scoring table: BLASTSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: Issued Parents AA:

1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/6C.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/6D.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 173.5 | 11.4 | 969 | 2 | US-08-284-941-2 |
| 2 | 173.5 | 11.4 | 969 | 2 | US-08-447-642-2 |
| 3 | 173.5 | 11.4 | 969 | 4 | US-09-236-503-2 |
| 4 | 173.5 | 11.4 | 969 | 5 | PCT-US93-02147A-2 |
| 5 | 162 | 10.7 | 799 | 2 | US-08-525-940-23 |
| 6 | 162 | 10.7 | 799 | 2 | US-08-976-838-23 |
| 7 | 162 | 10.7 | 881 | 2 | US-08-525-940-21 |
| 8 | 162 | 10.7 | 881 | 2 | US-08-976-838-21 |
| 9 | 162 | 10.7 | 915 | 2 | US-08-525-940-18 |
| 10 | 162 | 10.7 | 915 | 2 | US-08-976-838-18 |
| 11 | 162 | 10.7 | 915 | 4 | US-09-214-5558-2 |
| 12 | 162 | 10.7 | 915 | 4 | US-09-214-5558-7 |
| 13 | 161 | 10.6 | 288 | 1 | US-08-368-852-15 |
| 14 | 157.5 | 10.4 | 288 | 2 | US-08-525-940-15 |
| 15 | 157.5 | 10.4 | 288 | 1 | US-08-976-838-15 |
| 16 | 149 | 9.8 | 568 | 1 | US-07-862-021B-14 |
| 17 | 149 | 9.8 | 568 | 5 | PCT-US93-03164-14 |
| 18 | 148.5 | 9.8 | 2523 | 1 | US-08-185-432-18 |
| 19 | 148.5 | 9.8 | 2523 | 4 | US-08-899-232-18 |
| 20 | 147.5 | 9.7 | 802 | 1 | US-07-862-021B-12 |
| 21 | 147.5 | 9.7 | 802 | 1 | US-08-313-288B-12 |
| 22 | 147.5 | 9.7 | 802 | 5 | PCT-US93-03164-12 |
| 23 | 146 | 9.6 | 1068 | 1 | US-08-537-210A-2 |
| 24 | 146 | 9.6 | 1068 | 4 | US-09-113-825-2 |
| 25 | 146 | 9.6 | 2556 | 4 | US-08-185-432-17 |
| 26 | 146 | 9.6 | 2556 | 4 | US-08-899-232-2 |
| 27 | 143.5 | 9.5 | 807 | 1 | US-07-862-021B-10 |

| | | | | | | |
|----|-------|-----|------|---|-------------------|--------------------|
| 28 | 143.5 | 9.5 | 807 | 1 | US-08-313-288B-10 | Sequence 10, Appl |
| 29 | 143.5 | 9.5 | 807 | 5 | PCT-US93-03164-10 | Sequence 10, Appl |
| 30 | 139 | 9.2 | 366 | 4 | US-08-857-076-103 | Sequence 103, App |
| 31 | 139 | 9.2 | 486 | 3 | US-08-746-559A-5 | Sequence 5, Appl1 |
| 32 | 139 | 9.2 | 516 | 3 | US-08-746-559A-4 | Sequence 4, Appl1 |
| 33 | 139 | 9.2 | 1367 | 2 | US-08-249-687C-2 | Sequence 2, Appl1 |
| 34 | 139 | 9.2 | 1367 | 2 | US-08-625-819-2 | Sequence 2, Appl1 |
| 35 | 139 | 9.2 | 1367 | 3 | US-08-746-559A-2 | Sequence 2, Appl1 |
| 36 | 139 | 9.2 | 1367 | 4 | US-08-864-641B-18 | Sequence 18, Appl |
| 37 | 139 | 9.2 | 2556 | 4 | US-08-083-590A-20 | Sequence 20, Appl |
| 38 | 139 | 9.2 | 2556 | 3 | US-08-532-384-20 | Sequence 20, Appl |
| 39 | 135.5 | 8.9 | 370 | 4 | US-08-857-076-104 | Sequence 104, App |
| 40 | 135.5 | 8.9 | 1382 | 2 | US-08-737-715-2 | Sequence 2, Appl1 |
| 41 | 135.5 | 8.9 | 1382 | 4 | US-09-457-040B-7 | Sequence 2, Appl1 |
| 42 | 130 | 8.6 | 2471 | 1 | US-08-185-432-16 | Sequence 16, Appl1 |
| 43 | 130 | 8.6 | 2471 | 1 | US-08-083-590A-19 | Sequence 19, Appl |
| 44 | 130 | 8.6 | 2471 | 3 | US-08-532-384-19 | Sequence 19, Appl |
| 45 | 130 | 8.6 | 2471 | 4 | US-08-899-232-1 | Sequence 1, Appl1 |

ALIGNMENTS

```
RESULT 1
US-08-284-941-2
; Sequence 2, Application US/08284941
; Patent No. 5863756
;
GENERAL INFORMATION:
; APPLICANT: KIEFER, PHILIP J
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PAGE 4 AND
; TITLE OF INVENTION: PAGE 4.1 GENE AND POLYPEPTIDES IN CELLS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESS: COOLEY GODWARD CASTRO HUDDLESON & TATUM
; STREET: FIVE PALO ALTO SQUARE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94306
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,941
; FILING DATE: 2 August 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: NEELEY PH.D., RICHARD L.
; REGISTRATION NUMBER: 30092
; REFERENCE/DOCKET NUMBER: CHIR-009/01US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 843-5070
; TELEFAX: (415) 857-0663
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 969 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-284-941-2
;
Query Match 11.4%; Score 173.5; DB 2; Length 969;
Best Local Similarity 28.8%; Pred. No. 8e-07;
Matches 44; Conservative 16; Mismatches 54; Indels 39; Gaps 7;
QY 19 YIGSONASGRGRMRHNVNOCGCGCATCSD--YNGCLSKPRLFFALERIGMGIV 76
Db 737 YFEDTRARCR-----CHKGCCSSRAATQCLSCR-RGFY-----HHQEMNT 779
```


QY 77 CLSCPSGYYGTRYPDINKTKCKADCDTCFNK-NFTCKCKSGFYHLGKCLDNCPEGLE 135
DB 780 CVTLCPAGFYADE--SQKNCLCKHPCKCKVDEPEKCTVCKEGFSLARGSCIPDCEPGTY 837
QY 136 ANNHTECVSIVHCEVSEWNPWSPCTKKGKTCG 168
DB 838 FDSLLIRGCECHH-----TCG 853

RESULT 2
US-08-447-642-2
Sequence 2, Application US/08447642
Patent No. 5989890
GENERAL INFORMATION:
APPLICANT: BARR, PHILIP J
APPLICANT: KIEFER, MICHAEL C
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PAGE 4 AND
TITLE OF INVENTION: PAGE 4.1 GENE AND POLYPEPTIDES IN CELLS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOLEY GODDARD CASTRO HUDDLESON & TATUM
STREET: FIVE PALO ALTO SQUARE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,642
FILING DATE: 23-MAY-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/284,941
FILING DATE: 2 August 1994
ATTORNEY/AGENT INFORMATION:
NAME: NEELEY PH.D., RICHARD L.
REGISTRATION NUMBER: 30092
REFERENCE/DOCKET NUMBER: CHIR-009/01US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 843-5070
TELEFAX: (415) 857-0663
TELEX: 380816 COOLEY PA
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 969 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-447-642-2

Query Match 11.4%; Score 173.5; DB 2; Length 969;
Best Local Similarity 28.8%; Pred. No. 8e-07;
Matches 44; Conservative 16; Mismatches 54; Indels 39; Gaps 7;

QY 19 YIGSONASRGRRORHRHNPWSQCGGCATCSD--YNGCLSKPRLFFALRIGMKOIGV 76
DB 737 YFGDTARCR-----CHKCETCSSRAATGCLSR-RGFV-----HHQEWNT 779

QY 77 CLSCPSGYYGTRYPDINKTKCKADCDTCFNK-NFTCKCKSGFYHLGKCLDNCPEGLE 135
DB 780 CVTLCPAGFYADE--SQKNCLCKHPCKCKVDEPEKCTVCKEGFSLARGSCIPDCEPGTY 837

QY 136 ANNHTECVSIVHCEVSEWNPWSPCTKKGKTCG 168
DB 838 FDSLLIRGCECHH-----TCG 853

RESULT 3
US-09-236-503-2

Sequence 2, Application US/09236503
Patent No. 6277590
GENERAL INFORMATION:
APPLICANT: Barr, Philip J
APPLICANT: Kiefer, Michael C
TITLE OF INVENTION: Compositions and Methods for PAGE 4 and 4.1 Gene and
TITLE OF INVENTION: Polypeptides in Cells
FILE REFERENCE: CHIR-009/0405
CURRENT APPLICATION NUMBER: US/09/236,503
EARLIER FILING DATE: 1999-01-25
EARLIER APPLICATION NUMBER: 08/447,642
EARLIER FILING DATE: 1995-05-23
EARLIER APPLICATION NUMBER: 08/284,941
EARLIER FILING DATE: 1994-08-02
EARLIER APPLICATION NUMBER: 07/848,629
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 969
TYPE: PRT
ORGANISM: Homo sapiens
US-09-236-503-2

Query Match 11.4%; Score 173.5; DB 4; Length 969;
Best Local Similarity 28.8%; Pred. No. 8e-07;
Matches 44; Conservative 16; Mismatches 54; Indels 39; Gaps 7;

QY 19 YIGSONASRGRRORHRHNPWSQCGGCATCSD--YNGCLSKPRLFFALRIGMKOIGV 76
DB 737 YFGDTARCR-----CHKCETCSSRAATGCLSR-RGFV-----HHQEWNT 779

QY 77 CLSCPSGYYGTRYPDINKTKCKADCDTCFNK-NFTCKCKSGFYHLGKCLDNCPEGLE 135
DB 780 CVTLCPAGFYADE--SQKNCLCKHPCKCKVDEPEKCTVCKEGFSLARGSCIPDCEPGTY 837

QY 136 ANNHTECVSIVHCEVSEWNPWSPCTKKGKTCG 168
DB 838 FDSLLIRGCECHH-----TCG 853

RESULT 4
PCT-US93-02147A-2
Sequence 2, Application PC/TUS9302147A
GENERAL INFORMATION:
APPLICANT: BARR, PHILIP J
APPLICANT: KIEFER, MICHAEL C
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PAGE 4 AND
TITLE OF INVENTION: PAGE 4.1 GENE AND POLYPEPTIDES IN CELLS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOLEY GODDARD CASTRO HUDDLESON & TATUM
STREET: FIVE PALO ALTO SQUARE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/02147A
FILING DATE: 19930309
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,629
FILING DATE: 09-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: NEELEY PH.D., RICHARD L.
REGISTRATION NUMBER: 30092
REFERENCE/DOCKET NUMBER: CHIR-009/00US

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-7622
TELEFAX: (415) 857-0663
TELEX: 380816 COOLEY PA
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 969 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-02147A-2

Query Match 11.4%; Score 173.5; DB 5; Length 969;
Best Local Similarity 28.3%; Pred. No. 8e-07;
Matches 44; Conservative 16; Mismatches 54; Indels 39; Gaps 7;

Db 19 YIGSNASGRGRMRHNPVSGGCGCATCSD--YNGCLSCDPLFPALERIGMOICV 76
737 YGDDMARCR-----CHKQCTCSSRATQCLSCR-RGRY-----HROEMNT 779

Qy 77 CLSCDPSGYGTRYPDINKCTCKADCDTCFNNK-NFTCKKSGFYHLGKCLDNCPEGLE 135
780 CVTLPLPAFYADE--SOKNCLKCHPSCKKCVDEPEKCTVCEKGFSLARGSCIPDCEPGTY 837

Qy 136 ANNHTMECVSYHCEVSENNPWSPTCKGKTCG 168
838 PDESLIRGCECH-----TCG 853

Db 838 PDESLIRGCECH-----TCG 853

RESULT 5
US-08-525-940-23
Sequence 23, Application US/08525940
Patent No. 5866351
GENERAL INFORMATION:
APPLICANT: Franzusoff, Alex
APPLICANT: Miranda, Luis R.
APPLICANT: Wolf, Joseph R.
TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASES AND GENES
TITLE OF INVENTION: ENCODING SAID PROTEASES
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/525,940
APPLICATION NUMBER: US/08/525,940

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/368,852
FILING DATE: 01-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/088,322
FILING DATE: 07-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2848-11-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 799 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-525-940-23

Query Match 10.7%; Score 162; DB 2; Length 799;
Best Local Similarity 22.3%; Pred. No. 6.1e-06;
Matches 55; Conservative 33; Mismatches 81; Indels 78; Gaps 12;

Qy 21 GSONASGRGRMRHNPVSGGCGCATCSDYNG--CLSKPRLFPALERIGMOICV 78
563 GHYADK-KKCRKCAPN-----CESCFSGHODQMSCKGYFL-----NEETNSCV 607

Qy 79 SSCPSG-YGTRYPDINKCTCKADCDTCFNNK-NFTCKKSGFYHLGKCLDNCPEGLE 137
608 THCPDGSYODTKK--NLCKKSEKCTCTEFHNCCTECRDGLSLQSCRCSVSCEDRGYFN 664

Qy 138 NHTME-----CVS-----YHCEVSENNPWSPTCKK 163

Db 665 GQDQPCPFRFCATCAGAGDCTINCTEGYFMEDRCVOSCSISYTFHSSSENGYCKKC 724
164 GKTG-----GPKRGTE-----TRVEIIQHPAKGNLCPTNETRKTCTV 202

Db 725 DISCLTNGGPGFNCTSCPSGYLLDLGCMQGAICXDATEESWAEQFCMLVKKNNLC-- 782

Qy 203 GRKCKOK 209
783 QKRYLQ 789

Db 783 QKRYLQ 789

RESULT 6
US-08-976-838-23
Sequence 23, Application US/08976838
Patent No. 5981259
GENERAL INFORMATION:
APPLICANT: Franzusoff, Alex
TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASE NUCLEIC ACID
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln St., Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: US/08/976,838
APPLICATION NUMBER: US/08/976,838

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2848-11-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 799 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-976-838-23

Query Match 10.7%; Score 162; DB 2; Length 799;
Best Local Similarity 22.3%; Pred. No. 6.1e-06;
Matches 55; Conservative 33; Mismatches 81; Indels 78; Gaps 12;

QY 138 NHTME-----CVS-----IVHCEVSEW-----NPMSPCTKK 163
DB 747 GDDCPCHRFATCAGAGADGCTGCTGTFMEDRCVQSCSISYFPHSSSENGYKSCKK 806
QY 164 GKTG-----GFKRGTE-----TRVREIIQHPSAKGNLCPTNETRKTCTV 202
DB 807 DISCLTNGGPFKNCCTSCPSGYLLDLGMCWGAICXDATESWAEGBGFCMLVKKNMLC-- 864
QY 203 ORKKCOK 209
DB 865 QKKVLDQ 871

RESULT 9
US-08-525-940-18
Sequence 18, Application US/08525940
Patent No. 5866351
GENERAL INFORMATION:
APPLICANT: Franzusoff, Alex
APPLICANT: Miranda, Luis R.
TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASES AND GENES
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,940
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/368,852
FILING DATE: 01-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/088,322
FILING DATE: 07-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2848-11-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 915 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-525-940-18

Query Match 10.7%; Score 162; DB 2; Length 915;
Best Local Similarity 22.3%; Pred. No. 7,1e-06;
Matches 55; Conservative 33; Mismatches 81; Indels 78; Gaps 12;

QY 138 NHTME-----CVS-----IVHCEVSEW-----NPMSPCTKK 163
DB 761 GDDCPCHRFATCAGAGADGCTGCTGTFMEDRCVQSCSISYFPHSSSENGYKSCKK 840
QY 164 GKTG-----GFKRGTE-----TRVREIIQHPSAKGNLCPTNETRKTCTV 202
DB 841 DISCLTNGGPFKNCCTSCPSGYLLDLGMCWGAICXDATESWAEGBGFCMLVKKNMLC-- 898
QY 203 ORKKCOK 209
DB 899 QKKVLDQ 905

RESULT 10
US-08-976-838-18
Sequence 18, Application US/08976838
Patent No. 5981259
GENERAL INFORMATION:
APPLICANT: Franzusoff, Alex
TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASE NUCLEIC ACID
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln St., Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976,838
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2848-11-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 915 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-976-838-18

Query Match 10.7%; Score 162; DB 2; Length 915;
Best Local Similarity 22.3%; Pred. No. 7,1e-06;
Matches 55; Conservative 33; Mismatches 81; Indels 78; Gaps 12;

Qy 203 QRRKCK 209
Db 899 QRRKLOQ 905

RESULT 11

US-09-214-555B-2
Sequence 2, Application US/09214555B
Patent No. 6380171
GENERAL INFORMATION:
APPLICANT: INSTITUT DE RECHERCHE CLINIQUE DE MONTREAL
TITLE OF INVENTION: PRO-PROTEIN CONVERTING ENZYME
FILE REFERENCE: PRO-PROTEIN CONVERT ENZ
CURRENT APPLICATION NUMBER: US/09/214,555B
CURRENT FILING DATE: 1999-01-04
PRIOR APPLICATION NUMBER: 60/021,008
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: 2,203,745
PRIOR FILING DATE: 1997-04-25
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 915
TYPE: PR
ORGANISM: Homo sapiens
US-09-214-555B-2

Query Match 10.7%; Score 162; DB 4; Length 915;

Best Local Similarity 22.3%; Pred. No. 7,1e-06;
Matches 55; Conservative 33; Mismatches 81; Indels 78; Gaps 12;

Qy 21 GSONASRGRRORRMHPNVSGCGCATGSDYNG--CLSCPRLFALERIGMKQIGVCL 78
Db 679 GHYHADK-KRCKKAPN-----CESCFGSHDQCMCKGYFL-----NEETNSCV 723
Qy 79 SSCPSG-YGCTRPDIINCKTKKADCDTCFNKNKFCCKSGFYLLGKCLDNCEGLEAN 137
Db 724 THCPDSYDPTKK--NLCKKSENCKTCTEFHNCTECHRDLISLOGSRCSVSCEDGRYFN 780
Qy 138 NHTME-----CVS-----IVHCEVSEW-----NPMSPCTKK 163
Db 781 GDDCPCHRCATCAGAGADGICNCTEGYFMEDRCVQSCSISYFDHSSNGYKSCCKC 840
Qy 164 GKTC-----GFKRGTE-----TRVEIIOHPSAKGNLCPTNETRKCTV 202
Db 841 DISCLTNGPGFNCTSCPSGYLLDLGCMQMAICKDATEESMAEGFCMLVKKNLNC-- 898
Qy 203 QRRKCK 209
Db 899 QRRKLOQ 905

RESULT 12

US-09-214-555B-7
Sequence 7, Application US/09214555B
Patent No. 6380171
GENERAL INFORMATION:
APPLICANT: INSTITUT DE RECHERCHE CLINIQUE DE MONTREAL
TITLE OF INVENTION: PRO-PROTEIN CONVERTING ENZYME
FILE REFERENCE: PRO-PROTEIN CONVERT ENZ
CURRENT APPLICATION NUMBER: US/09/214,555B
CURRENT FILING DATE: 1999-01-04
PRIOR APPLICATION NUMBER: 60/021,008
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: 2,203,745
PRIOR FILING DATE: 1997-04-25
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 915
TYPE: PR
ORGANISM: Homo sapiens
US-09-214-555B-7

Query Match 10.7%; Score 162; DB 4; Length 915;
Best Local Similarity 22.3%; Pred. No. 7,1e-06;
Matches 55; Conservative 33; Mismatches 81; Indels 78; Gaps 12;

Qy 21 GSONASRGRRORRMHPNVSGCGCATGSDYNG--CLSCPRLFALERIGMKQIGVCL 78
Db 679 GHYHADK-KRCKKAPN-----CESCFGSHDQCMCKGYFL-----NEETNSCV 723
Qy 79 SSCPSG-YGCTRPDIINCKTKKADCDTCFNKNKFCCKSGFYLLGKCLDNCEGLEAN 137
Db 724 THCPDSYDPTKK--NLCKKSENCKTCTEFHNCTECHRDLISLOGSRCSVSCEDGRYFN 780
Qy 138 NHTME-----CVS-----IVHCEVSEW-----NPMSPCTKK 163
Db 781 GDDCPCHRCATCAGAGADGICNCTEGYFMEDRCVQSCSISYFDHSSNGYKSCCKC 840
Qy 164 GKTC-----GFKRGTE-----TRVEIIOHPSAKGNLCPTNETRKCTV 202
Db 841 DISCLTNGPGFNCTSCPSGYLLDLGCMQMAICKDATEESMAEGFCMLVKKNLNC-- 898
Qy 203 QRRKCK 209
Db 899 QRRKLOQ 905

RESULT 13

US-08-368-852-15
Sequence 15, Application US/08368852
Patent No. 5691183
GENERAL INFORMATION:
APPLICANT: Franzusoff, Alex
APPLICANT: Miranda, Luis R.
TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASES AND GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: CO
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/368,852
FILING DATE: 05-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: 2848-11
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/863-9700
TELEFAX: 303/863-0223
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 288 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-368-852-15

Query Match 10.6%; Score 161; DB 1; Length 288;
Best Local Similarity 28.2%; Pred. No. 2,4e-06;
Matches 46; Conservative 25; Mismatches 52; Indels 40; Gaps 11;

Qy 36 PNVSQ-GCQG-GCATSDYNGCLSCPRLFALERIGMKQIGVCLSSCPGYGTRYPDI 93
Db 899 QRRKLOQ 905

Db 13 PECSEVCGDGPDPHCHND---CL---HYYKXK---NNTRICVSSCPGHY---HADK 58
QY 94 NKCTCKKADCDTCTFNK--NFCTCKKSGFYH--LGKCLDNCEGIEANNHTMECVSIYHC 149
Db 59 KRCRCAPNCSCSCFGSHDQCMSCSKYGFLENETNSCVTHCPDSYODTKXNLG----- 112
QY 150 EVSEWNPMSPTCKKGTCTGFRGTETRV-REIIQHPASAGNLCPTNETRKTCTVORKKCO 191
Db 113 -----RKCSXNKCTCTEFTXKCTEFT-----DGLSLGSGSRC 142

RESULT 14
US-08-525-940-15
Sequence 15, Application US/08525940
Patent No. 5866351
GENERAL INFORMATION:
APPLICANT: Franzusoff, Alex
APPLICANT: Miranda, Luis R.
APPLICANT: Wolf, Joseph R.
TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASES AND GENES
TITLE OF INVENTION: ENCODING SAID PROTEASES
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,940
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/368,852
FILING DATE: 01-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/088,322
FILING DATE: 07-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2848-11-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 288 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-525-940-15

Query Match 10.4%; Score 157.5; DB 2; Length 288;
Best Local Similarity 26.6%; Pred. No. 4.7e-06;
Matches 49; Conservative 30; Mismatches 64; Indels 41; Gaps 12;

Db 13 PECSEVCGDGPDPHCHND---CL---HYYKXK---NNTRICVSSCPGHY---HADK 58
QY 94 NKCTCKKADCDTCTFNK--NFCTCKKSGFYH--LGKCLDNCEGIEANNHTMECVSIYHC 149
Db 59 KRCRCAPNCSCSCFGSHDQCMSCSKYGFLENETNSCVTHCPDSYODTKXNLG----- 112
QY 150 EVSEWNPMSPTCKKGTCTGFRGTETRV-REIIQHPASAGNLCPTNETRKTCTVORKKCO 208

Db 113 -----RKCSXNKCTCTEFTXKCTEFT-----DGLSLGSGSRC 158
QY 209 KGER 212
Db 159 PCHR 162

RESULT 15
US-08-976-838-15
Sequence 15, Application US/08976838
Patent No. 5981259
GENERAL INFORMATION:
APPLICANT: Franzusoff, Alex
APPLICANT: Miranda, Luis R.
APPLICANT: Wolf, Joseph R.
TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASE NUCLEIC ACID
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln St., Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976,838
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2848-11-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 288 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-976-838-15

Query Match 10.4%; Score 157.5; DB 2; Length 288;
Best Local Similarity 26.6%; Pred. No. 4.7e-06;
Matches 49; Conservative 30; Mismatches 64; Indels 41; Gaps 12;

Db 13 PECSEVCGDGPDPHCHND---CL---HYYKXK---NNTRICVSSCPGHY---HADK 58
QY 94 NKCTCKKADCDTCTFNK--NFCTCKKSGFYH--LGKCLDNCEGIEANNHTMECVSIYHC 149
Db 59 KRCRCAPNCSCSCFGSHDQCMSCSKYGFLENETNSCVTHCPDSYODTKXNLG----- 112
QY 150 EVSEWNPMSPTCKKGTCTGFRGTETRV-REIIQHPASAGNLCPTNETRKTCTVORKKCO 208
Db 113 -----RKCSXNKCTCTEFTXKCTEFT-----DGLSLGSGSRC 158

QY 209 KGER 212
Db 159 PCHR 162

Search completed: May 6, 2003, 14:53:24
Job time: 17 secs

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L11894; AAA61987.1; -
DR HSSP: O99405; 1MPT.
DR MEROPS: S08.075; -
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR002884; P domain.
DR InterPro: IPR002029; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR Pfam: PF01483; P; PARTIAL.
DR PRINTS: PR00723; SUBTILISIN.
DR PRODOM: PD000717; P domain; 1.
DR SMART: SM00281; EGF; 1.
DR SMART: SM00281; FU; 5.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
KW Hydrolyase; Serine protease; Glycoprotein; Zymogen; Signal; Calcium;
KW Cleavage on pair of basic residues; Repeat.
FT SIGNAL 1 45
FT PROPEP 46 132
FT CHAIN 133 937
FT FT
FT DOMAIN 133 454
FT FT 477 615
FT DOMAIN 680 937
FT SITE 132 133
FT SITE 534 536
FT ACT_SITE 186 166
FT ACT_SITE 227 227
FT ACT_SITE 401 401
FT CARBOHYD 240 401
FT CARBOHYD 882 882
FT CARBOHYD 900 900
SQ SEQUENCE 937 AA; 104053 MW; F3865557C33705C8 CRC64;
Query Match 10.3%; Score 156.5; DB 1; Length 937;
Best Local Similarity 23.6%; Pred. No. 0.0005; Mismatches 80; Indels 61; Gaps 11;
Matches 51; Conservative 24; Mismatches 80; Indels 61; Gaps 11;
QY 19 YIGQNASRGRORRMHNVSGCCGCGCATCSQDYV--GCLSCKRLFFALERIGMKQIGV 76
DB 705 YFGDTAARCR-----CHKGCECTGSRSPTOCLSCR-RGFY-----HNOETNT 747
QY 77 CLSSCPGVGTRYPDINKCKKADCDPCFNK-NFTCKKSGFYHLGKCLDNCPEGLE 135
DB 748 CVTLCPAGLVADDSQRL--CLRCHPSCKCKVDEPEKSTVCKEFSLARSGCIPDEPGTY 805
QY 136 ANHHTMEC-----VSTVHCEVS---EWMPSPTCKKKGKCTGFKGTETR 176
DB 806 FDSGLIFGCEHHTCTCVGSPSRRECIHCAKSFHPODMKCVPAQGE-----GF----- 853
QY 177 VREIIHPSAKGNLCPTNETRCKTQVRKCKQKGR 212
DB 854 -----YPEEMPGL--PHKYCRCDENCLSCGSGSR 881
RESULT 9
FSPQ_XENLA STANDARD; PRT; 803 AA.
AC P35447;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE F-spondin precursor.

OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8155;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=93376785; PubMed=8367492;
RA Ruiz i Altaba A., Cox C., Jessell T.M., Klar A.;
RT "Ectopic neural expression of a floor plate maker in frog embryos
RT injected with the midline transcription factor Pnt3a1aav.1";
RL Proc. Natl. Acad. Sci. U.S.A. 90:8268-8272(1993).
CC -1- FUNCTION: PROMOTES THE ATTACHMENT OF SPINAL CORD AND SENSORY
CC NEURON CELLS AND THE OUTGROWTH OF NEURITES IN VITRO. MAY
CC CONTRIBUTE TO THE GROWTH AND GUIDANCE OF AXONS IN BOTH THE SPINAL
CC CORD AND THE PNS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE FLOOR PLATE.
CC -1- SIMILARITY: CONTAINS 6 TSP TYPE-1 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L09123; AAA19105.1; -
DR PIR: A47723; A47723.
DR InterPro: IPR002861; Reeler.
DR InterPro: IPR00884; TSP1.
DR Pfam: PF00090; TSP 1; 6.
DR Pfam: PF02014; Reeler; 1.
DR SMART: SM00209; Reeler; 1.
DR PROSITE: PS50092; TSP1; 6.
KW Glycoprotein; Signal; Repeat; Cell adhesion.
FT SIGNAL 1 23
FT CHAIN 24 803
FT DOMAIN 437 488
FT DOMAIN 496 548
FT DOMAIN 553 604
FT DOMAIN 609 661
FT DOMAIN 665 716
FT DOMAIN 751 803
FT CARBOHYD 210 210
FT CARBOHYD 677 677
SQ SEQUENCE 803 AA; 90702 MW; D3A54E329548AED9 CRC64;
Query Match 10.3%; Score 156; DB 1; Length 803;
Best Local Similarity 24.6%; Pred. No. 0.00048; Mismatches 91; Indels 52; Gaps 12;
Matches 58; Conservative 35; Mismatches 91; Indels 52; Gaps 12;
QY 38 VSQCGCGCATSDVNGCLSKRLFFALERIGMKQIGVCLSSCP--GYGTRYPDINK 95
DB 544 VNECEPSSCIETWAEWECSS---ATCRMKKGRHMTMTADDSMKADTTVEK 598
QY 96 C--TKCK-----ADCDTCFNKNECTKCKSGFYHLGKCLDNCPEGLEANN-- 138
DB 599 CMAPCEHTIPCVLSFWSEWSDSVTCGKGTTROR---MLKPSFLGCGNTELELKQVEK 655
QY 139 -HMECVSTVHCEVSEWPMSPCTKKGKGTGFGRTGRTVREIIHPSAKGNLCPTNET 197
DB 656 CMLPECP--ISCELTENYSEC---NKSQG--KGMIRTRMTITNEPFGAVPEPVOR 708
QY 198 RCTVQKKCKQKGERKKKGRKKKPKNGESKEALPDKSLSSKEKPEORENKQ 253
DB 709 KKCRL--RKCKR---SSGNERRLK-----DAREKRESEKIKEDSDGO 747
RESULT 10
NTC2_MOUSE

ID NTC2 MOUSE STANDARD; PRT: 2470 AA.
 AC 035516; 060941; 006008;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neurogenic locus notch homolog protein 2 precursor (Notch 2) (Notch
 B)
 GN NOTCH2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Thymus;
 RA Hamada Y., Higuchi M., Tsujimoto Y.;
 RT "Complete amino acid sequence and multiform transcripts encoded by a
 RT single copy of mouse Notch2 gene";
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 316-1518 FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Embryo;
 RA Lardelli M., Lendahl U.;
 RT "Notch A and Notch B-two mouse Notch homologues coexpressed in a
 RT wide variety of tissues";
 RL Exp. Cell Res. 204:364-372(1993).
 RN [3]
 RP SEQUENCE OF 1765-2153 FROM N.A.
 RC MEDLINE=97075110; PubMed=8917536;
 RA Milner L.A., Bigas A., Kopan R., Brahm-Stein C., Bernstein I.D.,
 RA Martin D.I.;
 RT "Inhibition of granulocytic differentiation by mNotch1";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13014-13019(1996).
 RN [4]
 RP FUNCTION.
 RC MEDLINE=99396706; PubMed=10393120;
 RA Hamada Y., Kadokawa Y., Okabe M., Ikawa M., Coleman J.R.,
 RA Tsujimoto Y.;
 RT "Mutation in ankyrin repeats of the mouse Notch2 gene induces early
 RT embryonic lethality";
 RL Development 126:3415-3424(1999).
 RN [5]
 RP DEVELOPMENTAL STAGE, AND ALTERNATIVE SPLICING.
 RC MEDLINE=99333893; PubMed=7609614;
 RA Higuchi M., Kiyama H., Hayakawa T., Hamada Y., Tsujimoto Y.;
 RT "Differential expression of Notch1 and Notch2 in developing and adult
 RT mouse brain";
 RL Brain Res. Mol. Brain Res. 29:263-272(1995).
 RN [6]
 RP POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS OF MET-1699.
 RC MEDLINE=21523956; PubMed=1158718;
 RA Saxena M.T., Schroeter E.H., Mumm J.S., Kopan R.;
 RT "Murine notch homologs (N1-4) undergo presenilin-dependent
 RT proteolysis";
 RL J. Biol. Chem. 276:40268-40273(2001).
 RN [7]
 RP POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS OF MET-1699.
 RC MEDLINE=21373376; PubMed=11559941;
 RA Mizutani T., Taniguchi Y., Aoki T., Hashimoto N., Honjo T.;
 RT "Conservation of the biochemical mechanisms of signal transduction
 RT among mammalian Notch family members";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9026-9031(2001).
 CC -1- FUNCTION: Functions as a receptor for membrane-bound ligands
 CC Jagged2 and Delta to regulate cell-face determination.
 CC Upon ligand activation through the released notch intracellular
 CC domain (NICD) it forms a transcriptional activator complex with
 CC RBP-J kappa and activates genes of the enhancer of split locus.
 CC Affects the implementation of differentiation, proliferation and
 CC apoptotic programs (by similarity). May play an essential role in
 CC postimplantation development, probably in some aspect of cell
 CC specification and/or differentiation.
 CC -1- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-

CC terminal fragment N(EC) which are probably linked by disulfide
 CC bonds.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Following
 CC proteolytical processing NICD is translocated to the nucleus.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2, may be
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Expressed in the brain, liver, kidney,
 CC neuroepithelia, somites, optic vesicles and branchial arches, but
 CC not heart.
 CC -1- DEVELOPMENTAL STAGE: Expressed in the embryonic ventricular zone,
 CC the postnatal ependymal cells, and the choroid plexus throughout
 CC embryonic and postnatal development.
 CC -1- PTM: Synthesized in the endoplasmic reticulum as an inactive form
 CC which is proteolytically cleaved by a furin-like convertase in the
 CC trans-golgi network before it reaches the plasma membrane to yield
 CC an active, ligand-accessible form. Cleavage results in a C-
 CC terminal fragment N(TM) and a N-terminal fragment N(EC). Following
 CC ligand binding, it is cleaved by TNF-alpha converting enzyme
 CC (TACE) to yield a membrane-associated intermediate fragment called
 CC notch extracellular truncation (NEXT). This fragment is then
 CC cleaved by presenilin dependent gamma-secretase to release a
 CC notch-derived peptide containing the intracellular domain (NICD)
 CC from the membrane.
 CC -1- PTM: Phosphorylated.
 CC -1- SIMILARITY: BELONGS TO THE NOTCH FAMILY.
 CC -1- SIMILARITY: CONTAINS 34.5 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 LIN/NOTCH REPEATS.
 CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL, on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D32210; BAA22094.1; -
 CC EMBL: X68279; CAA8340.1; -
 CC EMBL: U1881; AAC52924.1; -
 CC HSSP: P16109, 1PSB.
 CC MED: MGI:97364, Notch2.
 CC InterPro: IPR002110; ANK.
 CC InterPro: IPR000152; Asx_hydroxyl.
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR00742; EGF_2.
 CC InterPro: IPR01881; EGF_Ca.
 CC InterPro: IPR001438; EGF_11.
 CC InterPro: IPR008800; Notch.
 CC Pfam: PF00008; EGF_35.
 CC Pfam: PF00023; ank; 6.
 CC Pfam: PF00066; notch; 2.
 CC PRINTS: PR00010; EGFBLDOD.
 CC PRINTS: PR01452; NOTCH.
 CC SMART: SM00248; ANK; 4.
 CC SMART: SM00179; EGF_CA; 22.
 CC SMART: SM00001; EGF_Like; 12.
 CC SMART: SM00004; NL_3.
 CC PROSITE: PSS0068; ANK_REPEAT; 4.
 CC PROSITE: PSS0297; ANK_REPEAT; 4.
 CC PROSITE: PSS0010; ASX_HYDROXYL; 22.
 CC PROSITE: PSS0022; EGF_1; 33.
 CC PROSITE: PSS0186; EGF_2; 27.
 CC PROSITE: PSS0187; EGF_CA; 22.
 CC Receptor: Transcription regulation; Activator; Differentiation;
 CC Transmembrane protein; Repeat; ANK repeat; EGF-like domain;
 CC Transmembrane; Glycoprotein; Signal; Phosphorylation;
 CC Alternative splicing.
 CC SIGNAL 1 25
 CC CHAIN 26 2470
 CC CHAIN 1666 2470
 CC CHAIN 1697 2470
 CC DOMAIN 26 1677
 CC
 CC POTENTIAL.
 CC NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 2.
 CC NOTCH EXTRACELLULAR TRUNCATION.
 CC NOTCH INTRACELLULAR DOMAIN.
 CC EXTRACELLULAR (POTENTIAL).


```

FT TRANSMEM 1678 1698 POTENTIAL.
FT DOMAIN 1699 2470 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 26 63 EGF-LIKE 1.
FT DOMAIN 64 102 EGF-LIKE 2.
FT DOMAIN 105 143 EGF-LIKE 3.
FT DOMAIN 144 180 EGF-LIKE 4.
FT DOMAIN 182 219 EGF-LIKE 5.
FT DOMAIN 221 256 EGF-LIKE 6 (INCOMPLETE).
FT DOMAIN 258 294 EGF-LIKE 7.
FT DOMAIN 296 334 EGF-LIKE 8.
FT DOMAIN 336 372 EGF-LIKE 9.
FT DOMAIN 373 411 EGF-LIKE 10.
FT DOMAIN 413 452 EGF-LIKE 11.
FT DOMAIN 454 490 EGF-LIKE 12.
FT DOMAIN 492 528 EGF-LIKE 13.
FT DOMAIN 530 566 EGF-LIKE 14.
FT DOMAIN 568 603 EGF-LIKE 15.
FT DOMAIN 605 641 EGF-LIKE 16.
FT DOMAIN 643 678 EGF-LIKE 17.
FT DOMAIN 680 716 EGF-LIKE 18.
FT DOMAIN 718 753 EGF-LIKE 19.
FT DOMAIN 755 791 EGF-LIKE 20.
FT DOMAIN 793 829 EGF-LIKE 21.
FT DOMAIN 831 869 EGF-LIKE 22.
FT DOMAIN 871 907 EGF-LIKE 23.
FT DOMAIN 909 945 EGF-LIKE 24.
FT DOMAIN 947 983 EGF-LIKE 25.
FT DOMAIN 985 1021 EGF-LIKE 26.
FT DOMAIN 1023 1059 EGF-LIKE 27.
FT DOMAIN 1061 1097 EGF-LIKE 28.
FT DOMAIN 1099 1145 EGF-LIKE 29.
FT DOMAIN 1147 1183 EGF-LIKE 30.
FT DOMAIN 1185 1221 EGF-LIKE 31.
FT DOMAIN 1223 1260 EGF-LIKE 32.
FT DOMAIN 1262 1300 EGF-LIKE 33.
FT DOMAIN 1302 1345 EGF-LIKE 34.
FT DOMAIN 1372 1410 EGF-LIKE 35.
FT DOMAIN 1418 1454 LIN/NOTCH 1.
FT DOMAIN 1501 1533 LIN/NOTCH 2.
FT DOMAIN 1535 1569 ANK 1.
FT DOMAIN 1571 1605 ANK 2.
FT DOMAIN 1607 1641 ANK 3.
FT DOMAIN 1643 1677 ANK 4.
FT DOMAIN 1679 1713 ANK 5.
FT DOMAIN 1715 1749 ANK 6.
FT DOMAIN 1751 1785 ANK 6.
FT DOMAIN 1787 1821 POLY-ALA.
FT DOMAIN 1823 1857 POLY-LEU.
FT DOMAIN 1859 1893 POLY-ALA.
FT DOMAIN 1895 1929 POLY-SER.
FT DOMAIN 1931 1965 POLY-GLY.
FT DOMAIN 1967 2001 BY SIMILARITY.
FT DOMAIN 2003 2037 BY SIMILARITY.
FT DOMAIN 2039 2073 BY SIMILARITY.
FT DOMAIN 2075 2109 BY SIMILARITY.
FT DOMAIN 2111 2145 BY SIMILARITY.
FT DOMAIN 2147 2181 BY SIMILARITY.
FT DOMAIN 2183 2217 BY SIMILARITY.
FT DOMAIN 2219 2253 BY SIMILARITY.
FT DOMAIN 2255 2289 BY SIMILARITY.
FT DOMAIN 2291 2325 BY SIMILARITY.
FT DOMAIN 2327 2361 BY SIMILARITY.
FT DOMAIN 2363 2397 BY SIMILARITY.
FT DOMAIN 2399 2433 BY SIMILARITY.
FT DOMAIN 2435 2469 BY SIMILARITY.
FT DOMAIN 2471 2505 BY SIMILARITY.
FT DOMAIN 2507 2541 BY SIMILARITY.
FT DOMAIN 2543 2577 BY SIMILARITY.
FT DOMAIN 2579 2613 BY SIMILARITY.
FT DOMAIN 2615 2649 BY SIMILARITY.
FT DOMAIN 2651 2685 BY SIMILARITY.
FT DOMAIN 2687 2721 BY SIMILARITY.
FT DOMAIN 2723 2757 BY SIMILARITY.
FT DOMAIN 2759 2793 BY SIMILARITY.
FT DOMAIN 2795 2829 BY SIMILARITY.
FT DOMAIN 2831 2865 BY SIMILARITY.
FT DOMAIN 2867 2901 BY SIMILARITY.
FT DOMAIN 2903 2937 BY SIMILARITY.
FT DOMAIN 2939 2973 BY SIMILARITY.
FT DOMAIN 2975 3009 BY SIMILARITY.
FT DOMAIN 3011 3045 BY SIMILARITY.
FT DOMAIN 3047 3081 BY SIMILARITY.
FT DOMAIN 3083 3117 BY SIMILARITY.
FT DOMAIN 3119 3153 BY SIMILARITY.
FT DOMAIN 3155 3189 BY SIMILARITY.
FT DOMAIN 3191 3225 BY SIMILARITY.
FT DOMAIN 3227 3261 BY SIMILARITY.
FT DOMAIN 3263 3297 BY SIMILARITY.
FT DOMAIN 3299 3333 BY SIMILARITY.
FT DOMAIN 3335 3369 BY SIMILARITY.
FT DOMAIN 3371 3405 BY SIMILARITY.
FT DOMAIN 3407 3441 BY SIMILARITY.
FT DOMAIN 3443 3477 BY SIMILARITY.
FT DOMAIN 3479 3513 BY SIMILARITY.
FT DOMAIN 3515 3549 BY SIMILARITY.
FT DOMAIN 3551 3585 BY SIMILARITY.
FT DOMAIN 3587 3621 BY SIMILARITY.
FT DOMAIN 3623 3657 BY SIMILARITY.
FT DOMAIN 3659 3693 BY SIMILARITY.
FT DOMAIN 3695 3729 BY SIMILARITY.
FT DOMAIN 3731 3765 BY SIMILARITY.
FT DOMAIN 3767 3801 BY SIMILARITY.
FT DOMAIN 3803 3837 BY SIMILARITY.
FT DOMAIN 3839 3873 BY SIMILARITY.
FT DOMAIN 3875 3909 BY SIMILARITY.
FT DOMAIN 3911 3945 BY SIMILARITY.
FT DOMAIN 3947 3981 BY SIMILARITY.
FT DOMAIN 3983 4017 BY SIMILARITY.
FT DOMAIN 4019 4053 BY SIMILARITY.
FT DOMAIN 4055 4089 BY SIMILARITY.
FT DOMAIN 4091 4125 BY SIMILARITY.
FT DOMAIN 4127 4161 BY SIMILARITY.
FT DOMAIN 4163 4197 BY SIMILARITY.
FT DOMAIN 4199 4233 BY SIMILARITY.
FT DOMAIN 4235 4269 BY SIMILARITY.
FT DOMAIN 4271 4305 BY SIMILARITY.
FT DOMAIN 4307 4341 BY SIMILARITY.
FT DOMAIN 4343 4377 BY SIMILARITY.
FT DOMAIN 4379 4413 BY SIMILARITY.
FT DOMAIN 4415 4449 BY SIMILARITY.
FT DOMAIN 4451 4485 BY SIMILARITY.
FT DOMAIN 4487 4521 BY SIMILARITY.
FT DOMAIN 4523 4557 BY SIMILARITY.
FT DOMAIN 4559 4593 BY SIMILARITY.
FT DOMAIN 4595 4629 BY SIMILARITY.
FT DOMAIN 4631 4665 BY SIMILARITY.
FT DOMAIN 4667 4701 BY SIMILARITY.
FT DOMAIN 4703 4737 BY SIMILARITY.
FT DOMAIN 4739 4773 BY SIMILARITY.
FT DOMAIN 4775 4809 BY SIMILARITY.
FT DOMAIN 4811 4845 BY SIMILARITY.
FT DOMAIN 4847 4881 BY SIMILARITY.
FT DOMAIN 4883 4917 BY SIMILARITY.
FT DOMAIN 4919 4953 BY SIMILARITY.
FT DOMAIN 4955 4989 BY SIMILARITY.
FT DOMAIN 4991 5025 BY SIMILARITY.
FT DOMAIN 5027 5061 BY SIMILARITY.
FT DOMAIN 5063 5097 BY SIMILARITY.
FT DOMAIN 5099 5133 BY SIMILARITY.
FT DOMAIN 5135 5169 BY SIMILARITY.
FT DOMAIN 5171 5205 BY SIMILARITY.
FT DOMAIN 5207 5241 BY SIMILARITY.
FT DOMAIN 5243 5277 BY SIMILARITY.
FT DOMAIN 5279 5313 BY SIMILARITY.
FT DOMAIN 5315 5349 BY SIMILARITY.
FT DOMAIN 5351 5385 BY SIMILARITY.
FT DOMAIN 5387 5421 BY SIMILARITY.
FT DOMAIN 5423 5457 BY SIMILARITY.
FT DOMAIN 5459 5493 BY SIMILARITY.
FT DOMAIN 5495 5529 BY SIMILARITY.
FT DOMAIN 5531 5565 BY SIMILARITY.
FT DOMAIN 5567 5601 BY SIMILARITY.
FT DOMAIN 5603 5637 BY SIMILARITY.
FT DOMAIN 5639 5673 BY SIMILARITY.
FT DOMAIN 5675 5709 BY SIMILARITY.
FT DOMAIN 5711 5745 BY SIMILARITY.
FT DOMAIN 5747 5781 BY SIMILARITY.
FT DOMAIN 5783 5817 BY SIMILARITY.
FT DOMAIN 5819 5853 BY SIMILARITY.
FT DOMAIN 5855 5889 BY SIMILARITY.
FT DOMAIN 5891 5925 BY SIMILARITY.
FT DOMAIN 5927 5961 BY SIMILARITY.
FT DOMAIN 5963 5997 BY SIMILARITY.
FT DOMAIN 5999 6033 BY SIMILARITY.
FT DOMAIN 6035 6069 BY SIMILARITY.
FT DOMAIN 6071 6105 BY SIMILARITY.
FT DOMAIN 6107 6141 BY SIMILARITY.
FT DOMAIN 6143 6177 BY SIMILARITY.
FT DOMAIN 6179 6213 BY SIMILARITY.
FT DOMAIN 6215 6249 BY SIMILARITY.
FT DOMAIN 6251 6285 BY SIMILARITY.
FT DOMAIN 6287 6321 BY SIMILARITY.
FT DOMAIN 6323 6357 BY SIMILARITY.
FT DOMAIN 6359 6393 BY SIMILARITY.
FT DOMAIN 6395 6429 BY SIMILARITY.
FT DOMAIN 6431 6465 BY SIMILARITY.
FT DOMAIN 6467 6501 BY SIMILARITY.
FT DOMAIN 6503 6537 BY SIMILARITY.
FT DOMAIN 6539 6573 BY SIMILARITY.
FT DOMAIN 6575 6609 BY SIMILARITY.
FT DOMAIN 6611 6645 BY SIMILARITY.
FT DOMAIN 6647 6681 BY SIMILARITY.
FT DOMAIN 6683 6717 BY SIMILARITY.
FT DOMAIN 6719 6753 BY SIMILARITY.
FT DOMAIN 6755 6789 BY SIMILARITY.
FT DOMAIN 6791 6825 BY SIMILARITY.
FT DOMAIN 6827 6861 BY SIMILARITY.
FT DOMAIN 6863 6897 BY SIMILARITY.
FT DOMAIN 6899 6933 BY SIMILARITY.
FT DOMAIN 6935 6969 BY SIMILARITY.
FT DOMAIN 6971 7005 BY SIMILARITY.
FT DOMAIN 7007 7041 BY SIMILARITY.
FT DOMAIN 7043 7077 BY SIMILARITY.
FT DOMAIN 7079 7113 BY SIMILARITY.
FT DOMAIN 7115 7149 BY SIMILARITY.
FT DOMAIN 7151 7185 BY SIMILARITY.
FT DOMAIN 7187 7221 BY SIMILARITY.
FT DOMAIN 7223 7257 BY SIMILARITY.
FT DOMAIN 7259 7293 BY SIMILARITY.
FT DOMAIN 7295 7329 BY SIMILARITY.
FT DOMAIN 7331 7365 BY SIMILARITY.
FT DOMAIN 7367 7401 BY SIMILARITY.
FT DOMAIN 7403 7437 BY SIMILARITY.
FT DOMAIN 7439 7473 BY SIMILARITY.
FT DOMAIN 7475 7509 BY SIMILARITY.
FT DOMAIN 7511 7545 BY SIMILARITY.
FT DOMAIN 7547 7581 BY SIMILARITY.
FT DOMAIN 7583 7617 BY SIMILARITY.
FT DOMAIN 7619 7653 BY SIMILARITY.
FT DOMAIN 7655 7689 BY SIMILARITY.
FT DOMAIN 7691 7725 BY SIMILARITY.
FT DOMAIN 7727 7761 BY SIMILARITY.
FT DOMAIN 7763 7797 BY SIMILARITY.
FT DOMAIN 7799 7833 BY SIMILARITY.
FT DOMAIN 7835 7869 BY SIMILARITY.
FT DOMAIN 7871 7905 BY SIMILARITY.
FT DOMAIN 7907 7941 BY SIMILARITY.
FT DOMAIN 7943 7977 BY SIMILARITY.
FT DOMAIN 7979 8013 BY SIMILARITY.
FT DOMAIN 8015 8049 BY SIMILARITY.
FT DOMAIN 8051 8085 BY SIMILARITY.
FT DOMAIN 8087 8121 BY SIMILARITY.
FT DOMAIN 8123 8157 BY SIMILARITY.
FT DOMAIN 8159 8193 BY SIMILARITY.
FT DOMAIN 8195 8229 BY SIMILARITY.
FT DOMAIN 8231 8265 BY SIMILARITY.
FT DOMAIN 8267 8301 BY SIMILARITY.
FT DOMAIN 8303 8337 BY SIMILARITY.
FT DOMAIN 8339 8373 BY SIMILARITY.
FT DOMAIN 8375 8409 BY SIMILARITY.
FT DOMAIN 8411 8445 BY SIMILARITY.
FT DOMAIN 8447 8481 BY SIMILARITY.
FT DOMAIN 8483 8517 BY SIMILARITY.
FT DOMAIN 8519 8553 BY SIMILARITY.
FT DOMAIN 8555 8589 BY SIMILARITY.
FT DOMAIN 8591 8625 BY SIMILARITY.
FT DOMAIN 8627 8661 BY SIMILARITY.
FT DOMAIN 8663 8697 BY SIMILARITY.
FT DOMAIN 8699 8733 BY SIMILARITY.
FT DOMAIN 8735 8769 BY SIMILARITY.
FT DOMAIN 8771 8805 BY SIMILARITY.
FT DOMAIN 8807 8841 BY SIMILARITY.
FT DOMAIN 8843 8877 BY SIMILARITY.
FT DOMAIN 8879 8913 BY SIMILARITY.
FT DOMAIN 8915 8949 BY SIMILARITY.
FT DOMAIN 8951 8985 BY SIMILARITY.
FT DOMAIN 8987 9021 BY SIMILARITY.
FT DOMAIN 9023 9057 BY SIMILARITY.
FT DOMAIN 9059 9093 BY SIMILARITY.
FT DOMAIN 9095 9129 BY SIMILARITY.
FT DOMAIN 9131 9165 BY SIMILARITY.
FT DOMAIN 9167 9201 BY SIMILARITY.
FT DOMAIN 9203 9237 BY SIMILARITY.
FT DOMAIN 9239 9273 BY SIMILARITY.
FT DOMAIN 9275 9309 BY SIMILARITY.
FT DOMAIN 9311 9345 BY SIMILARITY.
FT DOMAIN 9347 9381 BY SIMILARITY.
FT DOMAIN 9383 9417 BY SIMILARITY.
FT DOMAIN 9419 9453 BY SIMILARITY.
FT DOMAIN 9455 9489 BY SIMILARITY.
FT DOMAIN 9491 9525 BY SIMILARITY.
FT DOMAIN 9527 9561 BY SIMILARITY.
FT DOMAIN 9563 9597 BY SIMILARITY.
FT DOMAIN 9599 9633 BY SIMILARITY.
FT DOMAIN 9635 9669 BY SIMILARITY.
FT DOMAIN 9671 9705 BY SIMILARITY.
FT DOMAIN 9707 9741 BY SIMILARITY.
FT DOMAIN 9743 9777 BY SIMILARITY.
FT DOMAIN 9779 9813 BY SIMILARITY.
FT DOMAIN 9815 9849 BY SIMILARITY.
FT DOMAIN 9851 9885 BY SIMILARITY.
FT DOMAIN 9887 9921 BY SIMILARITY.
FT DOMAIN 9923 9957 BY SIMILARITY.
FT DOMAIN 9959 9993 BY SIMILARITY.
FT DOMAIN 9995 10000 BY SIMILARITY.

```

Query Match Best local Similarity 9.9%; Score 150.5; DB 1; Length 2470; Pred. No. 0.0031; Mismatches 53; Indels 89; Gaps 12;

```

DB 91 P-DINKCTKCKADDCITCFNNFC-----TKCKSGFYLIHLGK 125
DB 983 ENNIDECTE-----SSCFNGGTGTCVGINSFCLCPVGFTGPFCLHIDINCCSNPLCLNAGT 1037
DB 126 CLDN-----CEGLNANNHMECVSIYH-CEVSENNPSPSPTCKKGGKTCFRKGTETRV 177
DB 1038 CVDGLGTYRCTCPGLGYTKN-----CQTLVNLCS-----RSPKNGK-TC----- 1076
DB 178 REIIQHSAGKNLCPTNETRKTCTVQRKCC-----QKG 210

```

```

DB 1077 ---VQKARPHCLCPGMDGAYCDVLNVSCAAALQKG 1111
RESULT 11
ID IGIR_RAT STANDARD; PRT; 1370 AA.
AC P24062;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Insulin-like growth factor I receptor precursor (EC 2.7.1.112).
GN IGIR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxId=10116;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95277910; PubMed=7758167;
RA Du J., Delafontaine P.;
RT "Inhibition of vascular smooth muscle cell growth through antisense
transcription of a rat insulin-like growth factor I receptor cDNA.";
RC Cite. Res. 76:963-972(1995).
[2]
RP SEQUENCE OF 1-364 FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=90017496; PubMed=2477843;
RA Werner H., Moloschak W., Adamo M., Shen-Orr Z., Roberts C.T. Jr.,
RA Lerofth D.;
RT "Developmental regulation of the rat insulin-like growth factor I
receptor gene.";
RC Proc. Natl. Acad. Sci. U.S.A. 86:7451-7455(1989).
[3]
RP SEQUENCE OF 913-1017 FROM N.A.
RX MEDLINE=92412145; PubMed=1530648;
RA Kurachi H., Jodo K., Ohta M., Kawasaki T., Itoh N.;
RT "A new member of the insulin receptor family, insulin
receptor-related receptor, is expressed preferentially in the
kidney.";
RC Biochem. Biophys. Res. Commun. 187:934-939(1992).
CC - FUNCTION: THIS RECEPTOR BINDS INSULIN-LIKE GROWTH FACTOR I (IGF I)
WITH A HIGH AFFINITY AND IGF II WITH A LOWER AFFINITY. IT HAS A
TYROSINE-PROTEIN KINASE ACTIVITY.
CC - CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
CC - SUBUNIT: Tetramer of 2 alpha and 2 beta chains linked by disulfide
bonds. The alpha chains contribute to the formation of the ligand-
binding domain, while the beta chain carries the kinase domain.
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
RECEPTOR SUBFAMILY.
CC - SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL: L29232; AAA4192.1;
CC EMBL: M27293; AAA41384.1;
CC PIR: A33837; A33837.
CC HSSP: P06213; 11RK.
DR InterPro: IPR000494; EGF_R_L domain.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR002011; RTkinaseII.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00041; fn3; 2.

```


| | | | | |
|----|--------|-------|-------|---------------|
| FT | DOMAIN | 58 | 99 | EGF-LIKE 2. |
| FT | DOMAIN | 102 | 140 | EGF-LIKE 3. |
| FT | DOMAIN | 141 | 177 | EGF-LIKE 4. |
| FT | DOMAIN | 179 | 215 | EGF-LIKE 5. |
| FT | DOMAIN | 217 | 254 | EGF-LIKE 6. |
| FT | DOMAIN | 256 | 292 | EGF-LIKE 7. |
| FT | DOMAIN | 294 | 332 | EGF-LIKE 8. |
| FT | DOMAIN | 334 | 370 | EGF-LIKE 9. |
| FT | DOMAIN | 371 | 409 | EGF-LIKE 10. |
| FT | DOMAIN | 411 | 449 | EGF-LIKE 11. |
| FT | DOMAIN | 451 | 487 | EGF-LIKE 12. |
| FT | DOMAIN | 489 | 525 | EGF-LIKE 13. |
| FT | DOMAIN | 527 | 563 | EGF-LIKE 14. |
| FT | DOMAIN | 565 | 600 | EGF-LIKE 15. |
| FT | DOMAIN | 602 | 638 | EGF-LIKE 16. |
| FT | DOMAIN | 640 | 675 | EGF-LIKE 17. |
| FT | DOMAIN | 677 | 713 | EGF-LIKE 18. |
| FT | DOMAIN | 715 | 750 | EGF-LIKE 19. |
| FT | DOMAIN | 752 | 788 | EGF-LIKE 20. |
| FT | DOMAIN | 790 | 826 | EGF-LIKE 21. |
| FT | DOMAIN | 828 | 866 | EGF-LIKE 22. |
| FT | DOMAIN | 868 | 904 | EGF-LIKE 23. |
| FT | DOMAIN | 906 | 942 | EGF-LIKE 24. |
| FT | DOMAIN | 944 | 980 | EGF-LIKE 25. |
| FT | DOMAIN | 982 | 1018 | EGF-LIKE 26. |
| FT | DOMAIN | 1020 | 1056 | EGF-LIKE 27. |
| FT | DOMAIN | 1058 | 1094 | EGF-LIKE 28. |
| FT | DOMAIN | 1096 | 1142 | EGF-LIKE 29. |
| FT | DOMAIN | 1144 | 1180 | EGF-LIKE 30. |
| FT | DOMAIN | 1182 | 1218 | EGF-LIKE 31. |
| FT | DOMAIN | 1220 | 1264 | EGF-LIKE 32. |
| FT | DOMAIN | 1266 | 1304 | EGF-LIKE 33. |
| FT | DOMAIN | 1306 | 1346 | EGF-LIKE 34. |
| FT | DOMAIN | 1347 | 1383 | EGF-LIKE 35. |
| FT | DOMAIN | 1386 | 1424 | EGF-LIKE 36. |
| FT | REPEAT | 1441 | 1478 | EGF-LIKE 37. |
| FT | REPEAT | 1479 | 1520 | EGF-LIKE 38. |
| FT | REPEAT | 1521 | 1560 | EGF-LIKE 39. |
| FT | REPEAT | 1576 | 1619 | EGF-LIKE 40. |
| FT | REPEAT | 1624 | 1661 | EGF-LIKE 41. |
| FT | REPEAT | 1691 | 1737 | EGF-LIKE 42. |
| FT | REPEAT | 1737 | 1787 | EGF-LIKE 43. |
| FT | REPEAT | 1787 | 1837 | EGF-LIKE 44. |
| FT | REPEAT | 1837 | 1887 | EGF-LIKE 45. |
| FT | REPEAT | 1887 | 1937 | EGF-LIKE 46. |
| FT | REPEAT | 1937 | 1987 | EGF-LIKE 47. |
| FT | REPEAT | 1987 | 2037 | EGF-LIKE 48. |
| FT | REPEAT | 2037 | 2087 | EGF-LIKE 49. |
| FT | REPEAT | 2087 | 2137 | EGF-LIKE 50. |
| FT | REPEAT | 2137 | 2187 | EGF-LIKE 51. |
| FT | REPEAT | 2187 | 2237 | EGF-LIKE 52. |
| FT | REPEAT | 2237 | 2287 | EGF-LIKE 53. |
| FT | REPEAT | 2287 | 2337 | EGF-LIKE 54. |
| FT | REPEAT | 2337 | 2387 | EGF-LIKE 55. |
| FT | REPEAT | 2387 | 2437 | EGF-LIKE 56. |
| FT | REPEAT | 2437 | 2487 | EGF-LIKE 57. |
| FT | REPEAT | 2487 | 2537 | EGF-LIKE 58. |
| FT | REPEAT | 2537 | 2587 | EGF-LIKE 59. |
| FT | REPEAT | 2587 | 2637 | EGF-LIKE 60. |
| FT | REPEAT | 2637 | 2687 | EGF-LIKE 61. |
| FT | REPEAT | 2687 | 2737 | EGF-LIKE 62. |
| FT | REPEAT | 2737 | 2787 | EGF-LIKE 63. |
| FT | REPEAT | 2787 | 2837 | EGF-LIKE 64. |
| FT | REPEAT | 2837 | 2887 | EGF-LIKE 65. |
| FT | REPEAT | 2887 | 2937 | EGF-LIKE 66. |
| FT | REPEAT | 2937 | 2987 | EGF-LIKE 67. |
| FT | REPEAT | 2987 | 3037 | EGF-LIKE 68. |
| FT | REPEAT | 3037 | 3087 | EGF-LIKE 69. |
| FT | REPEAT | 3087 | 3137 | EGF-LIKE 70. |
| FT | REPEAT | 3137 | 3187 | EGF-LIKE 71. |
| FT | REPEAT | 3187 | 3237 | EGF-LIKE 72. |
| FT | REPEAT | 3237 | 3287 | EGF-LIKE 73. |
| FT | REPEAT | 3287 | 3337 | EGF-LIKE 74. |
| FT | REPEAT | 3337 | 3387 | EGF-LIKE 75. |
| FT | REPEAT | 3387 | 3437 | EGF-LIKE 76. |
| FT | REPEAT | 3437 | 3487 | EGF-LIKE 77. |
| FT | REPEAT | 3487 | 3537 | EGF-LIKE 78. |
| FT | REPEAT | 3537 | 3587 | EGF-LIKE 79. |
| FT | REPEAT | 3587 | 3637 | EGF-LIKE 80. |
| FT | REPEAT | 3637 | 3687 | EGF-LIKE 81. |
| FT | REPEAT | 3687 | 3737 | EGF-LIKE 82. |
| FT | REPEAT | 3737 | 3787 | EGF-LIKE 83. |
| FT | REPEAT | 3787 | 3837 | EGF-LIKE 84. |
| FT | REPEAT | 3837 | 3887 | EGF-LIKE 85. |
| FT | REPEAT | 3887 | 3937 | EGF-LIKE 86. |
| FT | REPEAT | 3937 | 3987 | EGF-LIKE 87. |
| FT | REPEAT | 3987 | 4037 | EGF-LIKE 88. |
| FT | REPEAT | 4037 | 4087 | EGF-LIKE 89. |
| FT | REPEAT | 4087 | 4137 | EGF-LIKE 90. |
| FT | REPEAT | 4137 | 4187 | EGF-LIKE 91. |
| FT | REPEAT | 4187 | 4237 | EGF-LIKE 92. |
| FT | REPEAT | 4237 | 4287 | EGF-LIKE 93. |
| FT | REPEAT | 4287 | 4337 | EGF-LIKE 94. |
| FT | REPEAT | 4337 | 4387 | EGF-LIKE 95. |
| FT | REPEAT | 4387 | 4437 | EGF-LIKE 96. |
| FT | REPEAT | 4437 | 4487 | EGF-LIKE 97. |
| FT | REPEAT | 4487 | 4537 | EGF-LIKE 98. |
| FT | REPEAT | 4537 | 4587 | EGF-LIKE 99. |
| FT | REPEAT | 4587 | 4637 | EGF-LIKE 100. |
| FT | REPEAT | 4637 | 4687 | EGF-LIKE 101. |
| FT | REPEAT | 4687 | 4737 | EGF-LIKE 102. |
| FT | REPEAT | 4737 | 4787 | EGF-LIKE 103. |
| FT | REPEAT | 4787 | 4837 | EGF-LIKE 104. |
| FT | REPEAT | 4837 | 4887 | EGF-LIKE 105. |
| FT | REPEAT | 4887 | 4937 | EGF-LIKE 106. |
| FT | REPEAT | 4937 | 4987 | EGF-LIKE 107. |
| FT | REPEAT | 4987 | 5037 | EGF-LIKE 108. |
| FT | REPEAT | 5037 | 5087 | EGF-LIKE 109. |
| FT | REPEAT | 5087 | 5137 | EGF-LIKE 110. |
| FT | REPEAT | 5137 | 5187 | EGF-LIKE 111. |
| FT | REPEAT | 5187 | 5237 | EGF-LIKE 112. |
| FT | REPEAT | 5237 | 5287 | EGF-LIKE 113. |
| FT | REPEAT | 5287 | 5337 | EGF-LIKE 114. |
| FT | REPEAT | 5337 | 5387 | EGF-LIKE 115. |
| FT | REPEAT | 5387 | 5437 | EGF-LIKE 116. |
| FT | REPEAT | 5437 | 5487 | EGF-LIKE 117. |
| FT | REPEAT | 5487 | 5537 | EGF-LIKE 118. |
| FT | REPEAT | 5537 | 5587 | EGF-LIKE 119. |
| FT | REPEAT | 5587 | 5637 | EGF-LIKE 120. |
| FT | REPEAT | 5637 | 5687 | EGF-LIKE 121. |
| FT | REPEAT | 5687 | 5737 | EGF-LIKE 122. |
| FT | REPEAT | 5737 | 5787 | EGF-LIKE 123. |
| FT | REPEAT | 5787 | 5837 | EGF-LIKE 124. |
| FT | REPEAT | 5837 | 5887 | EGF-LIKE 125. |
| FT | REPEAT | 5887 | 5937 | EGF-LIKE 126. |
| FT | REPEAT | 5937 | 5987 | EGF-LIKE 127. |
| FT | REPEAT | 5987 | 6037 | EGF-LIKE 128. |
| FT | REPEAT | 6037 | 6087 | EGF-LIKE 129. |
| FT | REPEAT | 6087 | 6137 | EGF-LIKE 130. |
| FT | REPEAT | 6137 | 6187 | EGF-LIKE 131. |
| FT | REPEAT | 6187 | 6237 | EGF-LIKE 132. |
| FT | REPEAT | 6237 | 6287 | EGF-LIKE 133. |
| FT | REPEAT | 6287 | 6337 | EGF-LIKE 134. |
| FT | REPEAT | 6337 | 6387 | EGF-LIKE 135. |
| FT | REPEAT | 6387 | 6437 | EGF-LIKE 136. |
| FT | REPEAT | 6437 | 6487 | EGF-LIKE 137. |
| FT | REPEAT | 6487 | 6537 | EGF-LIKE 138. |
| FT | REPEAT | 6537 | 6587 | EGF-LIKE 139. |
| FT | REPEAT | 6587 | 6637 | EGF-LIKE 140. |
| FT | REPEAT | 6637 | 6687 | EGF-LIKE 141. |
| FT | REPEAT | 6687 | 6737 | EGF-LIKE 142. |
| FT | REPEAT | 6737 | 6787 | EGF-LIKE 143. |
| FT | REPEAT | 6787 | 6837 | EGF-LIKE 144. |
| FT | REPEAT | 6837 | 6887 | EGF-LIKE 145. |
| FT | REPEAT | 6887 | 6937 | EGF-LIKE 146. |
| FT | REPEAT | 6937 | 6987 | EGF-LIKE 147. |
| FT | REPEAT | 6987 | 7037 | EGF-LIKE 148. |
| FT | REPEAT | 7037 | 7087 | EGF-LIKE 149. |
| FT | REPEAT | 7087 | 7137 | EGF-LIKE 150. |
| FT | REPEAT | 7137 | 7187 | EGF-LIKE 151. |
| FT | REPEAT | 7187 | 7237 | EGF-LIKE 152. |
| FT | REPEAT | 7237 | 7287 | EGF-LIKE 153. |
| FT | REPEAT | 7287 | 7337 | EGF-LIKE 154. |
| FT | REPEAT | 7337 | 7387 | EGF-LIKE 155. |
| FT | REPEAT | 7387 | 7437 | EGF-LIKE 156. |
| FT | REPEAT | 7437 | 7487 | EGF-LIKE 157. |
| FT | REPEAT | 7487 | 7537 | EGF-LIKE 158. |
| FT | REPEAT | 7537 | 7587 | EGF-LIKE 159. |
| FT | REPEAT | 7587 | 7637 | EGF-LIKE 160. |
| FT | REPEAT | 7637 | 7687 | EGF-LIKE 161. |
| FT | REPEAT | 7687 | 7737 | EGF-LIKE 162. |
| FT | REPEAT | 7737 | 7787 | EGF-LIKE 163. |
| FT | REPEAT | 7787 | 7837 | EGF-LIKE 164. |
| FT | REPEAT | 7837 | 7887 | EGF-LIKE 165. |
| FT | REPEAT | 7887 | 7937 | EGF-LIKE 166. |
| FT | REPEAT | 7937 | 7987 | EGF-LIKE 167. |
| FT | REPEAT | 7987 | 8037 | EGF-LIKE 168. |
| FT | REPEAT | 8037 | 8087 | EGF-LIKE 169. |
| FT | REPEAT | 8087 | 8137 | EGF-LIKE 170. |
| FT | REPEAT | 8137 | 8187 | EGF-LIKE 171. |
| FT | REPEAT | 8187 | 8237 | EGF-LIKE 172. |
| FT | REPEAT | 8237 | 8287 | EGF-LIKE 173. |
| FT | REPEAT | 8287 | 8337 | EGF-LIKE 174. |
| FT | REPEAT | 8337 | 8387 | EGF-LIKE 175. |
| FT | REPEAT | 8387 | 8437 | EGF-LIKE 176. |
| FT | REPEAT | 8437 | 8487 | EGF-LIKE 177. |
| FT | REPEAT | 8487 | 8537 | EGF-LIKE 178. |
| FT | REPEAT | 8537 | 8587 | EGF-LIKE 179. |
| FT | REPEAT | 8587 | 8637 | EGF-LIKE 180. |
| FT | REPEAT | 8637 | 8687 | EGF-LIKE 181. |
| FT | REPEAT | 8687 | 8737 | EGF-LIKE 182. |
| FT | REPEAT | 8737 | 8787 | EGF-LIKE 183. |
| FT | REPEAT | 8787 | 8837 | EGF-LIKE 184. |
| FT | REPEAT | 8837 | 8887 | EGF-LIKE 185. |
| FT | REPEAT | 8887 | 8937 | EGF-LIKE 186. |
| FT | REPEAT | 8937 | 8987 | EGF-LIKE 187. |
| FT | REPEAT | 8987 | 9037 | EGF-LIKE 188. |
| FT | REPEAT | 9037 | 9087 | EGF-LIKE 189. |
| FT | REPEAT | 9087 | 9137 | EGF-LIKE 190. |
| FT | REPEAT | 9137 | 9187 | EGF-LIKE 191. |
| FT | REPEAT | 9187 | 9237 | EGF-LIKE 192. |
| FT | REPEAT | 9237 | 9287 | EGF-LIKE 193. |
| FT | REPEAT | 9287 | 9337 | EGF-LIKE 194. |
| FT | REPEAT | 9337 | 9387 | EGF-LIKE 195. |
| FT | REPEAT | 9387 | 9437 | EGF-LIKE 196. |
| FT | REPEAT | 9437 | 9487 | EGF-LIKE 197. |
| FT | REPEAT | 9487 | 9537 | EGF-LIKE 198. |
| FT | REPEAT | 9537 | 9587 | EGF-LIKE 199. |
| FT | REPEAT | 9587 | 9637 | EGF-LIKE 200. |
| FT | REPEAT | 9637 | 9687 | EGF-LIKE 201. |
| FT | REPEAT | 9687 | 9737 | EGF-LIKE 202. |
| FT | REPEAT | 9737 | 9787 | EGF-LIKE 203. |
| FT | REPEAT | 9787 | 9837 | EGF-LIKE 204. |
| FT | REPEAT | 9837 | 9887 | EGF-LIKE 205. |
| FT | REPEAT | 9887 | 9937 | EGF-LIKE 206. |
| FT | REPEAT | 9937 | 9987 | EGF-LIKE 207. |
| FT | REPEAT | 9987 | 10037 | EGF-LIKE 208. |
| FT | REPEAT | 10037 | 10087 | EGF-LIKE 209. |
| FT | REPEAT | 10087 | 10137 | EGF-LIKE 210. |
| FT | REPEAT | 10137 | 10187 | EGF-LIKE 211. |
| FT | REPEAT | 10187 | 10237 | EGF-LIKE 212. |
| FT | REPEAT | 10237 | 10287 | EGF-LIKE 213. |
| FT | REPEAT | 10287 | 10337 | EGF-LIKE 214. |
| FT | REPEAT | 10337 | 10387 | EGF-LIKE 215. |
| FT | REPEAT | 10387 | 10437 | EGF-LIKE 216. |
| FT | REPEAT | 10437 | 10487 | EGF-LIKE 217. |
| FT | REPEAT | 10487 | 10537 | EGF-LIKE 218. |
| FT | REPEAT | 10537 | 10587 | EGF-LIKE 219. |
| FT | REPEAT | 10587 | 10637 | EGF-LIKE 220. |
| FT | REPEAT | 10637 | 10687 | EGF-LIKE 221. |
| FT | REPEAT | 10687 | 10737 | EGF-LIKE 222. |
| FT | REPEAT | 10737 | 10787 | EGF-LIKE 223. |
| FT | REPEAT | 10787 | 10837 | EGF-LIKE 224. |
| FT | REPEAT | 10837 | 10887 | EGF-LIKE 225. |
| FT | REPEAT | 10887 | 10937 | EGF-LIKE 226. |
| FT | REPEAT | 10937 | 10987 | EGF-LIKE 227. |
| FT | REPEAT | 10987 | 11037 | EGF-LIKE 228. |
| FT | REPEAT | 11037 | 11087 | EGF-LIKE 229. |
| FT | REPEAT | 11087 | 11137 | EGF-LIKE 230. |
| FT | REPEAT | 11137 | 11187 | EGF-LIKE 231. |
| FT | REPEAT | 11187 | 11237 | EGF-LIKE 232. |
| FT | REPEAT | 11237 | 11287 | EGF-LIKE 233. |
| FT | REPEAT | 11287 | 11337 | EGF-LIKE 234. |
| FT | REPEAT | 11337 | 11387 | EGF-LIKE 235. |
| FT | REPEAT | 11387 | 11437 | EGF-LIKE 236. |
| FT | REPEAT | 11437 | 11487 | EGF-LIKE 237. |
| FT | REPEAT | 11487 | 11537 | EGF-LIKE 238. |
| FT | REPEAT | 11537 | 11587 | EGF-LIKE 239. |
| FT | REPEAT | 11587 | 11637 | EGF-LIKE 240. |
| FT | REPEAT | 11637 | 11687 | EGF-LIKE 241. |
| FT | REPEAT | 11687 | 11737 | EGF-LIKE 242. |
| FT | REPEAT | 11737 | 11787 | EGF-LIKE 243. |
| FT | REPEAT | 11787 | 11837 | EGF-LIKE 244. |
| FT | REPEAT | 11837 | 11887 | EGF-LIKE 245. |
| FT | REPEAT | 11887 | 11937 | EGF-LIKE 246. |
| FT | REPEAT | 11937 | 11987 | EGF-LIKE 247. |
| FT | REPEAT | 11987 | 12037 | EGF-LIKE 248. |
| FT | REPEAT | 12037 | 12087 | EGF-LIKE 249. |
| FT | REPEAT | 12087 | 12137 | EGF-LIKE 250. |
| FT | REPEAT | 12137 | 12187 | EGF-LIKE 251. |
| FT | REPEAT | 12187 | 12237 | EGF-LIKE 252. |
| FT | REPEAT | 12237 | 12287 | EGF-LIKE 253. |
| FT | REPEAT | 12287 | 12337 | EGF-LIKE 254. |
| FT | REPEAT | 12337 | 12387 | EGF-LIKE 255. |
| FT | REPEAT | 12387 | 12437 | EGF-LIKE 256. |
| FT | REPEAT | 12437 | 12487 | EGF-LIKE 257. |
| FT | REPEAT | 12487 | 1 | |


```

Matches 54; Conservative 23; Mismatches 52; Indels 89; Gaps 13;

OY 31 QRMHNPNSGGCGGATGSDYNGCLSCKRPFLALERIGMIOIGVCLSSCPSSYIGTR 90
Db 947 QYDMNCSLSEPCNKG-GTCSDYNSYTC-----TCPPAFHCVHC 984
OY 91 P-DINKCTKCADCDTGFENKNC-----TKCKSGFYHLCK 125
Db 985 ENNIDECTE-----SSCFNGCTGVGINSFSCLPVGTGPFCLHDINECSNPCLNSGT 1039
OY 126 CLD-----NCEPGLEANNHMMCEVSLVH-CEVSEMNPMSCCTKKGTCGFKKGTETR 177
Db 1040 CVDGLCTKCTCPGTYGKN-----CQTLVNC-----SP-SCKKKG-TCA----- 1079
OY 178 RELIHPKAKGNLCPPNETRKYCTQRRKC-----QKG 210
Db 1080 ---CEKAPRCLCPGMDGAYCDVLVNSCKMAALQKG 1113

RESULT 14
BL14 CAEEL
ID BL14 CAEEL STANDARD: PRT: 943 AA.
AC PS159; 044762; 044763; 044764; 044765; 044766;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Endoprotease B11-4 precursor (EC 3.4.21.-) (Blisterase) (Blistered
DE cuticle protein 4).
GN B11-4 OR K04F10.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodirinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN RN
RP SEQUENCE FROM N.A. (ISOFORMS A; B; C AND D), AND FUNCTION.
RC STRAIN=Bristol N2;
RX MEDLINE=95293228; PubMed=7774813;
RA Thacker C., Peters K.W., Strayko M., Rose A.M.;
RT "The b11-4 locus of Caenorhabditis elegans encodes structurally
RT distinct kex2/subtilisin-like endoproteases essential for early
RT development and adult morphology.";
RL Gene Dev. 9:956-971(1995).
RN [2]
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RC STRAIN=Bristol N2;
RA Latreille P., Wamsley P.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: Isoform A is required for normal production or
CC maintenance of the adult cuticle. There is functional redundancy
CC between the isoforms. Vital for embryonic and larval development.
CC - TISSUE SPECIFICITY: In larvae and adults, expressed in all
CC hypodermal cells, vulva and ventral nerve cord.
CC - DEVELOPMENTAL STAGE: Expression starts at embryo two-fold stage
CC through to adults.
CC - ALTERNATIVE PRODUCTS: 5 isoforms: A/a, B/b, C/c, D/d (shown here)
CC and e, may be produced by alternative splicing.
CC - SIMILARITY: BELONGS TO REPETIDASE FAMILY S8, PRRIN SUBFAMILY.
CC - CAUTION: Rel.1 sequence differs from that shown due to a
CC frameshift in position 651.
CC - CAUTION: Ref.2 (AAB96754/AAB96757) sequence differs from that
CC shown due to erroneous gene model prediction.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use, by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL: L29438; AAA98750.1; ALT FRAME.
EMBL: L29439; AAA98751.1; ALT FRAME.
EMBL: L29440; AAA98752.1; ALT FRAME.

```

| AC | ID | WIFI_MOUSE | STANDARD; | PRT; | 379 AA. |
|----|--|-----------------------------------|-----------|------|--------------------------|
| AC | Q9WAL1 | WIFI_MOUSE | STANDARD; | PRT; | 379 AA. |
| DT | 16-OCT-2001 | (Rel. 40, Created) | | | |
| DT | 16-OCT-2001 | (Rel. 40, Last annotation update) | | | |
| DT | 16-OCT-2001 | (Rel. 40, Last sequence update) | | | |
| DE | Wnt inhibitory factor 1 precursor (WIF-1). | | | | |
| GN | WIF1. | | | | |
| OS | Mus musculus (Mouse). | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | | |
| OX | NCBI_TaxID=10090; | | | | |
| RN | [1]_TaxID=10090; | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RX | MEDLINE:99215557; PubMed:10201374; | | | | |
| RA | Heiter J.-C., Kodjabachian L., Rebder M.L., Ratner A., | | | | |
| RA | Shallwood P.M., Samos C.H., Nusse R., David I.B., Nathans J., | | | | |
| RT | "A new secreted protein that binds to Wnt proteins and inhibits their | | | | |
| RT | activities." | | | | |
| RL | Nature 398:431-436(1999). | | | | |
| CC | -1- FUNCTION: BINDS TO WNT PROTEINS AND INHIBITS THEIR ACTIVITIES. MAY | | | | |
| CC | BE INVOLVED IN MESODERM SEGMENTATION. | | | | |
| CC | -1- SUBCELLULAR LOCATION: Secreted. | | | | |
| CC | -1- TISSUE SPECIFICITY: EXPRESSION HIGHEST IN HEART AND LUNG. LOWER IN | | | | |
| CC | BRAIN AND EYE. | | | | |
| CC | -1- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS. | | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | | |
| CC | or send an email to license@isb-sib.ch). | | | | |
| CC | ----- | | | | |
| DR | EMBL; AF122923; AAD25403.1; - | | | | |
| DR | HSSP; P00740; IEDM. | | | | |
| DR | MGD; MG1:1344332; Wif1. | | | | |
| DR | InterPro; IPR000561; EGF-like | | | | |
| DR | InterPro; IPR002048; Laminin_EGF | | | | |
| DR | InterPro; IPR003306; WIF. | | | | |
| DR | Pfam; PF00008; EGF; 5. | | | | |
| DR | Pfam; PF02019; WIF; 1. | | | | |
| DR | PRINTS; PR00011; EGF/LAMININ. | | | | |
| DR | SMART; SM00181; EGF; 5. | | | | |
| DR | SMART; SM00469; WIF; 1. | | | | |
| DR | PROSITE; PS00022; EGF 1; 5. | | | | |
| DR | PROSITE; PS01186; EGF 2; 4. | | | | |
| KW | Repeat; EGF-like domain; signal; Developmental protein. | | | | |
| FT | CHAIN | 1 | 29 | 379 | POTENTIAL. |
| FT | DOMAIN | 177 | 208 | | WNT INHIBITORY FACTOR 1. |
| FT | DOMAIN | 209 | 240 | | EGF-LIKE 1. |
| FT | DOMAIN | 241 | 272 | | EGF-LIKE 2. |
| FT | DOMAIN | 273 | 304 | | EGF-LIKE 3. |
| FT | DOMAIN | 305 | 336 | | EGF-LIKE 4. |
| FT | DOMAIN | 337 | 368 | | EGF-LIKE 5. |
| FT | DISULFID | 177 | 186 | | POTENTIAL. |
| FT | DISULFID | 182 | 192 | | POTENTIAL. |
| FT | DISULFID | 198 | 200 | | POTENTIAL. |
| FT | DISULFID | 209 | 218 | | POTENTIAL. |
| FT | DISULFID | 214 | 224 | | POTENTIAL. |
| FT | DISULFID | 230 | 233 | | POTENTIAL. |
| FT | DISULFID | 241 | 250 | | POTENTIAL. |
| FT | DISULFID | 2 | | | |

FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL)
SQ SEQUENCE 379 AA; 41590 MM; E3765F2642B2BC9A CRC64;

Query Match 9.5%; Score 143.5; DB 1; Length 379;
Best Local Similarity 24.3%; Pred. No. 0.0019;
Matches 60; Conservative 24; Mismatches 68; Indels 95; Gaps 17;

QY 42 CCGGCAT---CSDYNGC-----LSCKPRLFFALERIGMKQIGVCLSS----C 81
DB 182 CPGGCRNGCFNERRVCECPDGFYGHCEKALCIPRCMNG-----GLCVTPGFCIC 232
QY 82 PGGYGTGYPDINKCTKC-KADCD-TCFNKNPCTCKSGFYHLGKCLDNCPEGLEANNH 139
DB 233 PPGFYG-----VNCDRANCSITCFNGGTC-----FY--PGKCI--CPGLEGE-- 271
QY 140 TMECVSIVHCEVSEMEFWSPTKKGKTCGFKRGTEYRVREIIQHPSA-KGNLCP----- 192
DB 272 -----QCELSKCP--QPCRNKGKCI-----KSKCKCPKGYGDLCKSKVCEP 312
QY 193 -----PTNETRKTCTVORKKCKOKGERGKKGRER-----KPKPKGSKKALIPDSKSL 240
DB 313 GCGAHGTCHPEPNC-----QCREGWHGRHCHNRRYGASLMHAPRPAGAGLERHTPPLKKA 367
QY 241 SSKELPE 247
DB 368 DRDPPE 374

Search completed: May 6, 2003, 14:52:00
Job time : 16 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:50:38 ; Search time 18 Seconds

(Without alignments) 1458.039 Million cell updates/sec

Title: US-09-894-912a-13

Perfect score: 1516

Sequence: 1 MGHRLISMFLINFMEMYI.....OQKRKVDKQSVSVTH 273

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 192 | 12.7 | 1299 | 2 T43251 | furin (EC 3.4.21.7 |
| 2 | 186.5 | 12.3 | 1548 | 2 S34583 | serine proteinase |
| 3 | 184 | 12.1 | 962 | 2 UC5571 | subtilisin-like pr |
| 4 | 184 | 12.1 | 975 | 2 UC5570 | subtilisin-like pr |
| 5 | 176.5 | 11.6 | 1680 | 2 A43434 | furin (EC 3.4.21.7 |
| 6 | 173.5 | 11.4 | 969 | 1 A39490 | subtilisin-like pr |
| 7 | 168 | 11.1 | 915 | 2 B48225 | probable propeptel |
| 8 | 167.5 | 11.0 | 932 | 2 I52527 | PACE4A - mouse (fr |
| 9 | 167 | 11.0 | 915 | 1 A48225 | subtilisin-like pr |
| 10 | 162 | 10.7 | 899 | 2 G02428 | subtilisin-like pr |
| 11 | 162 | 10.7 | 915 | 2 UC6148 | subtilisin-like pr |
| 12 | 161 | 10.6 | 440 | 2 T24232 | hypothetical prote |
| 13 | 158.5 | 10.5 | 379 | 2 A59180 | hypothetical prote |
| 14 | 156.5 | 10.3 | 937 | 2 I53282 | gene PAC4 protein |
| 15 | 156 | 10.3 | 803 | 2 A47723 | F-spondin precursor |
| 16 | 153.5 | 10.1 | 942 | 2 D87803 | protein bli-4D (lm |
| 17 | 150.5 | 9.9 | 1203 | 2 A49175 | protein B protein - |
| 18 | 149 | 9.8 | 1371 | 2 A33837 | insulin-like growt |
| 19 | 148.5 | 9.8 | 2524 | 2 A35844 | Xorch protein - Af |
| 20 | 147.5 | 9.7 | 2471 | 2 A49128 | cell-face determin |
| 21 | 145 | 9.6 | 570 | 2 T37314 | probable kexin (EC |
| 22 | 145 | 9.6 | 1620 | 2 T27283 | hypothetical prote |
| 23 | 143.5 | 9.5 | 807 | 2 A38152 | F-spondin - rat |
| 24 | 142 | 9.4 | 378 | 2 B59180 | Wnt inhibitory fac |
| 25 | 142 | 9.4 | 677 | 2 C42125 | trypophosite cystei |
| 26 | 139 | 9.2 | 1367 | 1 IGHUR1 | insulin-like growt |
| 27 | 139 | 9.2 | 2555 | 2 A50043 | notch protein homo |
| 28 | 138.5 | 9.1 | 1369 | 2 S70713 | protein-tyrosine k |
| 29 | 138 | 9.1 | 2531 | 2 S18188 | notch protein homo |

| | | | | | |
|----|-------|-----|------|----------|---------------------|
| 30 | 138 | 9.1 | 2531 | 2 A46019 | Notch-1 protein - |
| 31 | 137 | 9.0 | 2437 | 2 S42612 | transmembrane prot |
| 32 | 136 | 9.0 | 540 | 2 B47417 | insulin receptor-r |
| 33 | 135.5 | 8.9 | 1382 | 1 INHUR | insulin receptor p |
| 34 | 134.5 | 8.9 | 1111 | 2 T26972 | hypothetical prote |
| 35 | 134.5 | 8.9 | 1372 | 2 A34157 | insulin receptor p |
| 36 | 134.5 | 8.9 | 1383 | 2 A36080 | insulin receptor p |
| 37 | 129.5 | 8.5 | 2101 | 2 S57245 | insulin receptor (|
| 38 | 129.5 | 8.5 | 2148 | 1 A56081 | insulin receptor - |
| 39 | 128 | 8.4 | 327 | 2 A46484 | apoptosis-mediatin |
| 40 | 128 | 8.4 | 1268 | 2 B36502 | insulin receptor-r |
| 41 | 127 | 8.4 | 861 | 2 A48825 | Notch homolog Motc |
| 42 | 126.5 | 8.3 | 3871 | 2 T22812 | hypothetical prote |
| 43 | 125.5 | 8.3 | 667 | 2 A48579 | trypophosite surfac |
| 44 | 125 | 8.2 | 837 | 2 S43656 | furin (EC 3.4.21.7 |
| 45 | 124 | 8.2 | 1274 | 2 T42017 | cysteine rich prot |

ALIGNMENTS

RESULT 1
T43251
furin (EC 3.4.21.75) - fall armyworm
N:Alternate names: paired basic amino acid cleaving enzyme; propeptel convertase; seri
C:Species: Spodoptera frugiperda (fall armyworm)
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C/Accession: T43251
A:Description: Cloning and functional characterization of FURIN from Spodoptera frugipe
A:Reference number: Z22368
A:Accession: T43251
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1299 <CF>
A:Cross-references: EMBL:Z68888; NID:g1167859; PID:e2196590; PIDN:CAA91116.1
A:Experimental source: clone Sfurin 6; ovary
C:Function:
A:Description: responsible for the endoproteolytic processing of propeptides with speci
C/Keywords: hydrolase; serine proteinase

Query Match 12.7% Score 192; DB 2; Length 1299;
Best local similarity 27.8%; Pred No. 2.4e-05;
Matches 63; Conservative 26; Mismatches 78; Indels 60; Gaps 12;

OY 37 NVSGGCGGCGATCSQ-YNGLSCCKPRLPALERIGMKOIGVCLSSCPGYYGTRPDINK 95
Db 795 SVGRPCAHACITCSEBADGCTSCENHL-----VLHDTGTCMASCPESHYET---EDM 843
OY 96 CTRKADDPCTCF-NKQPCTRCKSGFYHLGKCLDNCPEGLAANNHTMECVSIHCEVSE 153
Db 844 CAKCHESCDTQCGEGTCTCHPSTYALDGCIVTSCPPAYADKKKKEC---MRCPVG- 899
OY 154 KNPWSPCTKRGKGTGFKRGTRVREIIQHPAKGNLCPP---TNETRKC-TVORKKCK 209
Db 900 ---CSTCT-----SAFLSCPEWELNKKKGMVGSQKCSA 933
OY 210 GERGKKRERKKKKPNK-----GESKE---AIPDSKLSSEKPEI 247
Db 934 GERAV---DQCKRCNPACDSCVGENEGHCLTCEPNNLDQYKCPVE 977

RESULT 2

S34583
serine proteinase (EC 3.4.21.-) PC6B - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C/Accession: S34583
R:Nakagawa, T.; Murakami, K.; Nakayama, K.
FEBS Lett. 327, 165-171, 1993
A>Title: Identification of an isoform with an extremely large Cys-rich region of PC6, a
A:Reference number: S34583; M0ID:93327934; PMID:835106

A:Accession: S34583
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1548 <NAK>
 A:Cross-references: GB:D17583; MID:9407344; PIDN:BAA04507.1; PID:d1005033; PID:g440374
 C:Keywords: hydrolase; serine proteinase

Query Match 12.3%; Score 186.5; DB 2; Length 1548;
 Best Local Similarity 25.0%; Pred. No. 6.3e-05;
 Matches 66; Conservative 31; Mismatches 90; Indels 77; Gaps 15;

39 SOGCGGCGATGSDYNG--CLSCKRLFFALERIGMKQIGVCLSCPGYGTGTRPDI--N 94
 DB 880 NQPCSSSKTC---NGSLCASCTGMYLMO-----ACVPSCTPG-----TWSEVTS 924
 QY 95 KTKTKKADCTGCPKXNPTCKCKS---GTYHLGKCLDNCPEGLANNHTME-CYSTVHC 149
 DB 925 SCEKSEDCVSCGADLCQOCLSPDNTLLHEGRCHYSCPEGYAADVCEHCSS--PC 982
 QY 150 EVSEWNPWSPCTKKG-----KTC-----GFKRGTRTRVEIIOHPSAKNL- 190
 DB 983 KTCCEGNATSCNSCGDFVLHDGVCMKTCPEKHVAVEGVCKHCPRCQDCIHEKTCCKMP 1042
 QY 191 -----CPPT--NETRKTCTQVRKKCKGKRGKGRKRRKPKNGSKKXKAIPLPS 236
 DB 1043 DFFLYNDMCHRSCKPSFYPMRQCVPCHKNC-----LECNQPKEDCKVCADTS 1091
 QY 237 KSLESS---KEIPE---ORENQ 253
 DB 1092 KALHNGLCIDCEPGETYKEEENDE 1115

RESULT 3
 JCS571
 subtilisin-like proprotein convertase (EC 3.4.21.-) PACE4 precursor, splice form E-I -
 C:Species: Homo sapiens (man)
 C:Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 20-Jun-2000
 C:Accession: JCS571
 R:Morit, K.; Kii, S.; Tsuji, A.; Nagahama, M.; Imamaki, A.; Hayashi, K.; Akamatsu, T.; Na
 J. Biochem. 121, 941-948, 1997
 A:Title: A novel human PACE4 isoform, PACE4E is an active processing protease containing
 A:Reference number: JCS570; MID:97335942; PMID:9192737
 A:Accession: JCS571
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-962 <MOR>
 A:Cross-references: DDBJ:D87994; MID:92330550; PIDN:BAA21791.1; PID:92330551
 A:Experimental source: brain cerebellum
 C:Comment: This enzyme is a processing protease and responsible for processing of vario
 ch it is retained intracellularly.
 C:Genetics:
 A:Gene: GDB:PACE4
 A:Cross-references: GDB:131390; OMIM:167405
 A:Map position: 15q26-15q26
 C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
 C:Keywords: glycoprotein; hydrolase; serine proteinase
 F:1-62/Domain: signal sequence #status predicted <SIG>
 F:63-149/Domain: propeptide #status predicted <PRO>
 F:196-434/Domain: subtilisin homology <SBT>
 F:938-954/Domain: hydrophobic cluster #status predicted <HCL>
 F:205-246,347-420/Active site: Asp, His, Asn, Ser #status predicted
 F:259/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 12.1%; Score 184; DB 2; Length 962;
 Best Local Similarity 26.6%; Pred. No. 6.4e-05;
 Matches 62; Conservative 24; Mismatches 87; Indels 60; Gaps 14;

19 YIGSONASGRGRORRHPNVSOQCGGCGATCSD--YNGCLSKPRLFFALERIGMKQIGV 76
 DB 724 YFGDPAARCR-----CHKCETCSSRAATCLSCR-RGFY-----HHQEMNT 766
 QY 77 CLSSPGSYGTRYPDINKCTCKACDCTCFNK-NFCTKCKSGFYHLGKCLDNCPEGLE 135
 DB 780 CVTLCPAGRYADE--SQNKCLKHPSCCKKVDPEPKCTVCKGFLANGSCIPCEGTY 837
 QY 136 ANNHTEEC-----VSIVHCEVS-ENPWSPTCKKCTGFRGTRVRE 179
 DB 838 FDELLRCECHHTGTCVGGRECHICKAKNFHDM-----KCVPAQGEPTPE----- 888
 QY 180 IIOHPSAKNLG-----PPTNETRKTCTQVRKKCKGKRGKGRKRRKPKNGE 228
 DB 889 --EMPGLPHKVCRRYRPGGE-RQATVS-----SKGVPG--GQSLASSPAGGE 932

DB 767 CVTLCPAGRYADE--SQNKCLKHPSCCKKVDPEPKCTVCKGFLANGSCIPCEGTY 824
 QY 136 ANNHTEEC-----VSIVHCEVS-ENPWSPTCKKCTGFRGTRVRE 179
 DB 825 FDELLRCECHHTGTCVGGRECHICKAKNFHDM-----KCVPAQGEPTPE----- 875
 QY 180 IIOHPSAKNLG-----PPTNETRKTCTQVRKKCKGKRGKGRKRRKPKNGE 228
 DB 876 --EMPGLPHKVCRRYRPGGE-RQATVS-----SKGVPG--GQSLASSPAGGE 919

RESULT 4
 JCS570
 subtilisin-like proprotein convertase (EC 3.4.21.-) PACE4 precursor, splice form E-I -
 C:Species: Homo sapiens (man)
 C:Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 20-Jun-2000
 C:Accession: JCS570
 R:Morit, K.; Kii, S.; Tsuji, A.; Nagahama, M.; Imamaki, A.; Hayashi, K.; Akamatsu, T.; N
 J. Biochem. 121, 941-948, 1997
 A:Title: A novel human PACE4 isoform, PACE4E is an active processing protease containin
 A:Reference number: JCS570; MID:97335942; PMID:9192737
 A:Accession: JCS570
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-975 <MOR>
 A:Cross-references: DDBJ:D87993; MID:92330548; PIDN:BAA21791.1; PID:92330549
 A:Experimental source: brain cerebellum
 C:Comment: This enzyme is a processing protease and responsible for processing of vario
 ch it is retained intracellularly.
 C:Genetics:
 A:Gene: GDB:PACE4
 A:Cross-references: GDB:131390; OMIM:167405
 A:Map position: 15q26-15q26
 C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
 C:Keywords: alternative splicing; glycoprotein; hydrolase; serine proteinase
 F:1-62/Domain: signal sequence #status predicted <SIG>
 F:63-149/Domain: propeptide #status predicted <PRO>
 F:196-434/Domain: subtilisin homology <SBT>
 F:938-954/Domain: hydrophobic cluster #status predicted <HCL>
 F:205-246,347-420/Active site: Asp, His, Asn, Ser #status predicted
 F:259/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 12.1%; Score 184; DB 2; Length 975;
 Best Local Similarity 26.6%; Pred. No. 6.5e-05;
 Matches 62; Conservative 24; Mismatches 87; Indels 60; Gaps 14;

19 YIGSONASGRGRORRHPNVSOQCGGCGATCSD--YNGCLSKPRLFFALERIGMKQIGV 76
 DB 737 YFGDPAARCR-----CHKCETCSSRAATCLSCR-RGFY-----HHQEMNT 779
 QY 77 CLSSPGSYGTRYPDINKCTCKACDCTCFNK-NFCTKCKSGFYHLGKCLDNCPEGLE 135
 DB 780 CVTLCPAGRYADE--SQNKCLKHPSCCKKVDPEPKCTVCKGFLANGSCIPCEGTY 837
 QY 136 ANNHTEEC-----VSIVHCEVS-ENPWSPTCKKCTGFRGTRVRE 179
 DB 838 FDELLRCECHHTGTCVGGRECHICKAKNFHDM-----KCVPAQGEPTPE----- 888
 QY 180 IIOHPSAKNLG-----PPTNETRKTCTQVRKKCKGKRGKGRKRRKPKNGE 228
 DB 889 --EMPGLPHKVCRRYRPGGE-RQATVS-----SKGVPG--GQSLASSPAGGE 932

RESULT 5
 A33434
 furin (EC 3.4.21.75) 2 - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 31-Mar-2000
 C:Accession: A43434
 J:Roobroeck, A.J.; Creemers, J.W.; Pauli, I.G.; Kurzik-Dumke, U.; Rentrop, M.; Gateff, E
 J. Biol. Chem. 267, 17208-17215, 1992
 A:Title: Cloning and functional expression of Dfurin2, a subtilisin-like proprotein prc
 A:Reference number: A43434; MID:92301056; PMID:1512255

A:Accession: A43434
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1680 <ROF>
 A:Cross-references: GB:M94375; NID:9157461; PID:9157462
 A>Note: sequence extracted from NCBI backbone (NCBIN:111933, NCBIP:111934)
 C:Genetics:
 A:Gene: FLYbase:Fur2
 A:Cross-references: FLYbase:FBgn0004598
 C:Superfamily: subtilisin homology
 C:Keywords: hydrolase; serine proteinase; transmembrane protein
 F:409-652/Domain: subtilisin homology <SBT>
 F:148-457,638/Active site: Asp, His, Ser #status predicted

Query Match 11.6%; Score 176.5; DB 2; Length 1680;
 Best Local Similarity 28.0%; Pred. No. 0.0003;
 Matches 60; Conservative 24; Mismatches 77; Indels 53; Gaps 12;

Db 3 HLRLSMFLILNFMFYGISGNASRGRRMRHNPVSGCGGATGSDY--NGCLSKRPR 61
 1038 HLHVID-LAVLQFCPPGYFENS-----RNRICVP-----CEPNASGODHPEYCTGCDH 1087
 Qy 62 LEFALERIGMKQIVCLSCPSGYGTRYPDINKCKKADCTCF--NKNFTCKKSGF 119
 1088 LVMHEHK-----CYSACPDTYET--EDNKAFCSTCATCNGPTDQCITCRSSR 1136
 Qy 120 YLHLCKCLDNCEGLBANHTMECVSIHACEVSENNPWSPTCKKGTGFGKGTETRY-R 178
 1137 YAMQKCKLISCPDGFADKRLKLCM-----PQEQCKTC-----TSNGVCS 1177
 Qy 179 ELIHPKSKNMLCPPTNTRKCTVQOKK-CQKGE 211
 1178 ECLQMT-----LNKRDKCIVSGSGGSESE 1203

RESULT 6
 A39490
 subtilisin-like proprotein convertase (EC 3.4.21.-) PACE4 precursor, splice form A - hum
 N:Alternate names: kexin homolog
 C:Species: Homo sapiens (man)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 31-Mar-2000
 C:Accession: A39490
 R:Kiefer, M.C.; Tucker, J.E.; Joh, R.; Landsberg, K.E.; Saltman, D.; Barr, P.J.;
 DNA Cell Biol. 10, 757-769, 1991
 A:Title: Identification of a second human subtilisin-like protease gene in the fes/fps r
 A:Reference number: A39490; MUID:92075167; PMID:1741956
 A:Accession: A39490
 A:Molecule type: mRNA
 A:Residues: 1-969 <KIE>
 A:Cross-references: GB:M80482; NID:9189531; PIDN:AAA59998.1; PID:9189532
 C:Genetics:
 A:Gene: GDB:PACE4
 A:Cross-references: GDB:131390; OMIM:167405
 A:Map position: 15q26-15q28
 C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
 C:Keywords: alternative splicing; hydrolase; serine proteinase
 F:150-969/Product: serine proteinase PACE4 #status predicted <SIG>
 F:196-444/Domain: subtilisin homology <SBT>
 F:205,246/Active site: Asp, His, Ser #status predicted

Query Match 11.4%; Score 173.5; DB 1; Length 969;
 Best Local Similarity 28.8%; Pred. No. 0.00032;
 Matches 44; Conservative 16; Mismatches 54; Indels 39; Gaps 7;

Db 19 YIGSONASRGRRMRHNPVSGCGGATGSDY--YNGCLSKRPLFALERIGMKQIV 76
 737 YFGDTAARRCR-----CHKGETSSRAATQCLSCR-RGTY-----HHQENMT 779
 Qy 77 CLSCPSGYGTRYPDINKCKKADCTCFNK-NFTCKKSGFYHLKCLDNCEGLE 135
 780 CVTLCPAGFADE--SQKXCKLCHPSCKKCVDEPKCTVCKEGFSIARSGCIPDCEPTY 837
 Qy 136 ANHTMECVSIHACEVSENNPWSPTCKKGTG 168

Db 838 FDSLEIRGECCH-----TCG 853

RESULT 7
 B48225
 probable proprotein convertase (EC 3.4.21.-) 5 precursor - rat
 N:Alternate names: PCS precursor
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 23-Feb-1997
 C:Accession: B48225
 R:Russon, U.; Vileau, D.; Hamelin, J.; Day, R.; Chretien, M.; Seidah, N.G.
 Proc. Natl. Acad. Sci. U.S.A. 90, 6691-6695, 1993
 A:Title: cDNA structure of the mouse and rat subtilisin/kexin-like PCS: a candidate pr
 A:Reference number: A48225; MUID:93342056; PMID:8341687
 A:Accession: B48225
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-915 <LUS>
 A:Cross-references: GB:U14933
 C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
 C:Keywords: duplication; glycoprotein; hydrolase; integrin binding; serine proteinase
 F:1-34/Domain: signal sequence #status predicted <SIG>
 F:35-915/Product: probable proprotein convertase 5 #status predicted <PRO>
 F:117-915/Product: probable proprotein convertase 5 #status experimental <MAT>
 F:164-402/Domain: subtilisin homology <SBT>
 F:173,214,388/Active site: Asp, His, Ser #status predicted

Query Match 11.1%; Score 168; DB 2; Length 915;
 Best Local Similarity 26.3%; Pred. No. 0.0007;
 Matches 49; Conservative 25; Mismatches 54; Indels 58; Gaps 11;

Db 36 PNVSQ-GCGG---GCATGSDYNGCLSKRPLFALERIGMKQIVCLSCPSGYGTRY 90
 640 PECSVCGDGPGRDCTDCLHYHKLGNTR-----ICVSGCPGHF---H 682
 Qy 91 PDINKCKKADCTCFNKNF--CTCKSGFYHL--LGKCLDNCEGLEANNHTMECVSI 146
 683 ADKKRRKCAPKPCSCPSHADQCLSKGYFLNERTSCVAGCPGSGYQIKKNTC--- 739
 Db 147 VHCVESENNPWSPTCKKKTCT-GFKGTETRYREIIGHPSAKNMLCPPTNTRKCTVQOK 205
 740 -----CKGSENKCTCTGFHNCTE-----CKGGL---SLGSRCSV--- 771
 Qy 206 KCKGE 211
 772 TCEDGQ 777

RESULT 8
 I52527
 PACE4 - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 16-Jul-1999
 C:Accession: I52527
 R:Hosaka, M.; Murakami, K.; Nakayama, K.
 Biomed. Res. 15, 383-390, 1994
 A:Title: PACE4A is a ubiquitous endoprotease that has similar but not identical substra
 A:Reference number: I52527
 A:Accession: I52527
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-932 <RES>
 A:Cross-references: GB:D50060; NID:9769700; PIDN:BA08777.1; PID:9769701
 C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
 F:112-810/Domain: subtilisin homology <SBT>

Query Match 11.0%; Score 167.5; DB 2; Length 932;
 Best Local Similarity 24.1%; Pred. No. 0.00076;
 Matches 52; Conservative 25; Mismatches 78; Indels 61; Gaps 11;

Db 19 YIGSONASRGRRMRHNPVSGCGGATGSDY--GCLSKRPLFALERIGMKQIV 76

RESULT 9
A48225
subtilisin-like proprotein convertase (EC 3.4.21.-) PCS precursor - mouse
N:Alternate names: kexin homology; serine proteinase PC6
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 Sequence_Revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: A48225; JX0248
R:Lusson, J.; Vleau, D.; Hamelin, J.; Day, R.; Chretien, M.; Seidah, N.G.
Proc. Natl. Acad. Sci. U.S.A. 90, 6691-6695, 1993
A:Title: CDNA structure of the mouse and rat subtilisin/kexin-like PCS: a candidate pro-
A:Reference number: A48225; MUID:93342056; PMID:831687
A:Accession: A48225
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-915 <LUS>
A:Cross-references: GB:114932; NID:9293327; PIDW:AA74636.1; PID:9293328
R:Nakagawa, T.; Hosaka, M.; Torii, S.; Matanabe, T.; Nakakami, K.; Nakayama, K.
J. Biochem. 113, 132-135, 1993
A:Title: Identification and functional expression of a new member of the mammalian Kex2
A:Reference number: JX0248; MUID:93324489; PMID:8468318
A:Accession: JX0248
A:Molecule type: mRNA
A:Residues: 1-915 <NAK>
A:Cross-references: DDBJ:DJ12619; NID:9220565; PIDN:BA02143.1; PID:9220566
A:Note: the authors translated the codon GGC for residue 915 as Ala
C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
C:Keywords: duplication; glycoprotein; hydrolysis; integrin binding; serine proteinase
F:1-34/Domain: signal sequence #status predicted <SIG>
F:35-116/Domain: propeptide #status predicted <PRO>
F:117-915/Product: proprotein convertase PCS #status experimental <MAT>
F:146-402/Domain: subtilisin homology <SBT>
F:173-214,388/Active site: Asp, His, Ser #status predicted

```
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #next_change 17-Mar-2000
C:Accession: G02428
R:Reudelhuber, T.L.
submitted to the EMBL Data Library, February 1996
A:Reference number: H01242
C:Accession: G02428
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-899 <REU>
A:Cross-references: EMBL:U49114, NID:G1218057, PIDD:AAA1807.1; PID:G121805
C:Genetics:
C:Gene: PC5
C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
C:Keywords: hydrolase; serine proteinase
F:148-386/Domain: subtilisin homology <Str>
```

```

Query Match      10.7%; Score 162; DB 2; Length 899;
Best Local Similarity 22.3%; Pred. No. 0.0017;
Matches 55; Conservative 33; Mismatches 81; Indels 78; Gaps 12;

OY 21 GSONASRGRRORRHPNVSGCGGCGATGSDYNG--CLSCKRRLPFALERIGMKOIGVCL 78
Db 663 GHVHADK-KRCKKAPN-----CESCGSHGDCMSCKKGYFL-----NEETNSCV 707
OY 79 SSCPSG-YYGTVRPDINKCTCKKADCDTCFNCNFKCTCKSGFYLHGKCLDNCPEGLEAN 137
Db 708 THRPDSYQDTKK---NLCKRCSENCCKTCTEFNINTECHDGLSLGSGCSVSCEDGRYFN 764
OY 138 NHTME-----CVS-----IVHCEVSEW-----NPSPTCKK 163
Db 765 GDDCPCHRFCATCAGAGADGCLNCTEGYFMEEDGRCVSGSISYVFDHSSSENGYSCKKC 824
OY 164 GKTC-----GFKRQTE-----TRVEIIIOHPSAQNLCPTNETRKCTV 202
Db 825 DISCLNCGPGRKNTCTSPSGYLLDLGMCQMGAIICDAVEESMAEGFCMLVKKNLIC-- 882
OY 203 QKKCKCK 209
Db 883 QKKVLDQ 889

RESULT 11
JC6148
subtilisin-like proprotein convertase (EC 3.4.21.-) homolog - human
N.Alternate names: PC6A protease
C.Species: Homo sapiens (man)
C.Date: 11-Apr-1997 #sequence, revision 09-May-1997 #text_change 22-Jun-1999
C.Accession: J06148
R.Miranda, L.; Wolf, J.; Pichuanes, S.; Duke, R.; Franzusoff, A.
Proc. Natl. Acad. Sci. U.S.A. 93, 7695-7700, 1996
A.Title: Isolation of the human PC6 gene encoding the putative host protease for HIV-1.
A.Reference number: J06148; MUID:96353880; PMID:8755538
A.Contents: CEM T-cell1
A.Accession: J06148
A.Molecule type: mRNA
A.Residues: 1-915 <MR>
A.Cross-references: GB:U56387; NID:g1498312; PIDN:AA050643.1; PID:g1498313
C.Comments: This protein functions as a soluble enzyme within the secretory pathway.
C.Genetics:
A:Gene: pc6a
C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
C:Keywords: glycoprotein; hydrolase; serine proteinase
F:164-402/Domain: subtilisin homology <SBT>

Query Match      10.7%; Score 162; DB 2; Length 915;
Best Local Similarity 22.3%; Pred. No. 0.0017;
Matches 55; Conservative 33; Mismatches 81; Indels 78; Gaps 12;

OY 21 GSONASRGRRORRHPNVSGCGGCGATGSDYNG--CLSCKRRLPFALERIGMKOIGVCL 78
Db 679 GHVHADK-KRCKKAPN-----CESCGSHGDCMSCKKGYFL-----NEETNSCV 723
OY 79 SSCPSG-YYGTVRPDINKCTCKKADCDTCFNCNFKCTCKSGFYLHGKCLDNCPEGLEAN 137
Db 883 QKKVLDQ 889

```

```

Db 724 THCPDSYVDTKX---NLCRCSCENCKTCTEPHNCCEGDLGLGSCSCSCEDEGRYFN 780
Qy 138 NHTME-----CVS-----YHCEVSEV-----NPMSPCTKK 163
Db 781 GDDCQCHRFCAATCAGAGADGCLNCTEGYFMEGRVCVSCSIYFDHSSENGKSKC 840
Qy 164 GKTC-----GFKRGTE-----TRVEIIOHPSAKGNLCPTNETRKTCTY 202
Db 841 DISCLTNGPGRFNCTSPSGYLLDLCMGMAICMDATSESWAGGFCMLVKKNNLC-- 898
Qy 203 QRRKKCK 209
Db 899 QRRVLDQ 905

```

RESULT 12

```

T24232
hypothetical protein R17.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T24232
R:Barlow, K.
submitted to the EMBL Data Library, March 1997
A:Reference number: Z19860
A:Accession: T24232
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-440 <M1>
A:Cross-references: EMBL:Z92809; PIDN:CAM07269.1; GSPDB:GN00021; CESP:R17.3
A:Experimental source: clone R17
C:Genetics:
A:Gene: CESP:R17.3
A:Map position: 3
A:Insertions: 36/3; 89/3; 179/2; 241/1; 306/1; 346/2
C:Superfamily: Caenorhabditis elegans hypothetical protein R17.3

```

```

Query Match 10.6%; Score 161; DB 2; Length 440;
Best Local Similarity 23.7%; Pred. No. 0.0012;
Matches 53; Conservative 30; Mismatches 87; Indels 54; Gaps 8;

Qy 59 KPRLPFALERIGMKOIGVCLSSCPGYYGTRYPDINKCTCKADCTCFNNKCTCKSG 118
Db 172 KPRHLIRYSLSKFMPKVTSTPLYENRVOFANNLYLESISCEYCDHC----- 225
Qy 119 FYLHIGKCLDNPEGLENNHTMECVSIHCEVSENNPSPCTKKGCTGCTETRRV 178
Db 226 --VTLLDCC-----CSDTFVCP-RDCLVLDMSHTQCTADNGTCG--GTQGRKR 271
Qy 179 EIIQHPSAKGNLCPTNETRKTCTVQRKKCGKGRGKGRERKKPKNGESKEAIPDSKS 238
Db 272 HVIQHAERGAACEPLKEMRTCFVE--CR-----PKKSALDD--- 306
Qy 239 LSSSKETPEORENKOOKKRR-----YDDKQSVSVSTVH 273
Db 307 ITTVALIDYRNKTRSKIRNNIYMDLPVNAEKMKKATYYCVH 350

```

RESULT 13

```

A59180
Wnt inhibitory factor-1 - human
C:Species: Homo sapiens (man)
C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
C:Accession: A59180
R:Heichen, J.C.; Kozljabachian, L.; Rebber, M.L.; Ratner, A.; Smallwood, P.M.; Samos, C.H.
Nature 398, 431-438, 1999
A>Title: A new secreted protein that binds to Wnt proteins and inhibits their activities
A:Reference number: A59180; MUID:99215557; PMID:10201374
A:Accession: A59180
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-379 <HS1>
A:Cross-references: GB:AF122922; NID:94585369; PIDN:AD25402.1; PID:94585370

```

```

Query Match 10.5%; Score 158.5; DB 2; Length 379;
Best Local Similarity 24.9%; Pred. No. 0.0015;
Matches 60; Conservative 26; Mismatches 72; Indels 83; Gaps 16;

Qy 42 CQGGCAT--CSDYNC-----LSCKPRLPFALERIGMKOIGVCLSS--C 81
Db 182 CPGGCRNNGFCNBRICECPDPFHGPHCEKALCTPRCMG-----GLCVTFPGCIC 232
Qy 82 PSGYYGTRYDINKCTKC-KADCD--TCFNKRFCTKCKSGFYHLGKCLDNCEGLEANNH 139
Db 233 PGRFYG-----VNCBKANCSTTCNNGTC-----FY-PGKCI--CPGLEGE-- 271
Qy 140 TMECVSIHCEVSENNPSPCTKGTGCFRGTETRVREIIQHPSAKGNLCPP----- 193
Db 272 -----OCEISKCP--QPCRNNGKCIQ--KSKCKSKGYQGLCSKPCVCEPGCGAHG 318
Qy 194 -TNETRKTCTVQRKKCGKGRGKGRERKX-----KPNKESKEAIPDSKLSSESKETP 246
Db 319 TCHERNKC-----QCEGHHGHCNKKRYEASLIHALRPAQAOIRQHTPLSKAEERRRPP 373
Qy 247 E 247
Db 374 E 374

```

RESULT 14

```

I53282
gene PACB4 protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 22-Jun-1999
C:Accession: I53282
R:Johnson, R.C.; Darlington, D.N.; Hand, T.A.; Bloomquist, B.T.; Mains, R.E.
Endocrinology 135, 1178-1185, 1994
A>Title: PACB4: a subtilisin-like endoprotease prevalent in the anterior pituitary and
A:Reference number: I53282; MUID:94349873; PMID:8070361
A:Accession: I53282
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-937 <RES>
A:Cross-references: GB:U13184; NID:9496221; PIDN:AAA61987.1; PID:9496222
C:Genetics:
A:Gene: PACB4
C:Superfamily: subtilisin-like proteinase PACB4; subtilisin homology
F.177-415/Domain: subtilisin homology <SBT>

```

```

Query Match 10.3%; Score 156.5; DB 2; Length 937;
Best Local Similarity 23.6%; Pred. No. 0.004;
Matches 51; Conservative 24; Mismatches 80; Indels 61; Gaps 11;

Qy 19 YIGSQNASRGRORRMHNPVSOQCGGCATCSDYN--GLSCKPRLPFALERIGMKOIGV 76
Db 705 YFGDTAARCR-----CHKGETCTGRSPQCLSCR-RGFY-----HHQETNT 747
Qy 77 CLSSCPGYYGTYPPDINKCTCKADCDTCFNNK-NFCTKCKSGFYHLGKCLDNCEGLE 135
Db 748 CVTLCPAGLYADESORL--CLKCHPSCKCVDEPKSTVCKESFLARSCCLPDEPETTY 805
Qy 136 ANNTMEC-----VSIVHCEVS-----EWNPMSPCTKKGCTGCFKRGTETR 176
Db 806 FDSRLIRGCECHHTCTCTCVGPERECICAKSFHFQDMKVPACGE-----GF----- 853
Qy 177 VREIIQHPSAKGNLCPTNETRKTCTVQRKKCGKGR 212
Db 854 -----YPEMPQL--PKVYCRDNDENCLSCGSSR 881

```

RESULT 15

```

A47723
F-spondin precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change 17-Nov-2000
C:Accession: A47723

```

R;Ruiz i Altaba, A.; Cox, C.; Jessell, T.M.; Klar, A.

Proc. Natl. Acad. Sci. U.S.A. 90, 8268-8272, 1993

A;Title: Ectopic neural expression of a floor plate marker in frog embryos injected with

A;Reference number: A47723; MUID:93376785; PMID:8367492

A;Accession: A47723

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-803 <RU1>

A;Cross-References: G8:L09123; NID:g409244; PIDN:AAA19105.1; PID:g409245

C;Superfamily: F-spondin; thrombospondin type 1 repeat homology

F;435-489/Domain: thrombospondin type 1 repeat homology <THR2>

F;607-662/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match

Best Local Similarity 10.3%; Score 156; DB 2; Length 803;

Matches 58; Conservative 35; Mismatches 91; Indels 52; Gaps 12;

```

Oy 38 VSOGCGGCGATCSDYNGCLSCPRLPFALERIGMQIGVCLSSCP--GYGTRYPDINK 95
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 544 VNECEPSSCIYTEMWEECS-----ATCRMGKKRRHRIMTPADGSMCKADTTEVER 598
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 96 C--TKCK-----ADCDTCFNQNECTKCKSGFYHLGKCLDNCPEGLEANN--- 138
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 599 CMWPECHTIPCVLSPMSFMSDCSVTCCKGTRTRQR---MLKSPSELGDCNELELKQVEK 655
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 139 -HMECVSIWCEVSEANPWSPECTKKGKTCGFKGTETREIIOHPSAKGNLCPTINET 197
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 656 CMLPECP--ISCLETWYSWSEC---NKSCG--KGHMRTRMTWEPQFGAVCPETVQR 708
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 198 RKCTVORKKCKGKRGKRRKKRKKPKNGESKEAIPDSKSLSSKEIPEQRENKQ 253
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 709 KRCRL--RKCCK---SSGNRRHLK-----DAREKRRREKIKEDSDGEQ 747
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

Search completed: May 6, 2003, 14:53:03

Job time : 20 secs

GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:49:17 ; Search time 37 seconds
(without alignments)
983.173 Million cell updates/sec

Title: US-09-894-912a-13

Perfect score: 1516
Sequence: 1 MGLRLISWLFILNFMET.....QKKRRVQKKQSVSVTVH 273

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: /SID52/gcgdata/geneeq/geneeq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneeq/geneeq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneeq/geneeq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneeq/geneeq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneeq/geneeq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneeq/geneeq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneeq/geneeq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneeq/geneeq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneeq/geneeq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneeq/geneeq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneeq/geneeq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneeq/geneeq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneeq/geneeq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneeq/geneeq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneeq/geneeq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneeq/geneeq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneeq/geneeq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneeq/geneeq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneeq/geneeq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneeq/geneeq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneeq/geneeq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneeq/geneeq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneeq/geneeq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Length | DB ID | Description |
|------------|-------|--------------|-------|-----------------------------|
| 1 | 1516 | 100.0 | 273 | AAE13151 Human stem cell gr |
| 2 | 1505 | 99.3 | 272 | AAE13150 Human stem cell gr |
| 3 | 1505 | 99.3 | 272 | AAE13168 Human stem cell gr |
| 4 | 1505 | 99.3 | 272 | AAW78328 Human protein SEQ |
| 5 | 1505 | 99.3 | 272 | AAW78320 Human thymospond |
| 6 | 1472 | 97.1 | 265 | AAE13163 Human secreted pro |
| 7 | 1472 | 97.1 | 292 | AAW85607 Secreted protein c |
| 8 | 1472 | 97.1 | 292 | AAE13170 Human SCR-1 relate |
| 9 | 1472 | 97.1 | 292 | ABP61846 Human polypeptide |
| 10 | 1452 | 95.8 | 292 | AAW93875 Human protein sequ |

| | | | | |
|----|--------|------|------|-----------------------------|
| 11 | 1400 | 92.3 | 251 | AAE13153 Human mature stem |
| 12 | 1310.5 | 86.4 | 279 | AAE13167 Mouse stem cell gr |
| 13 | 903 | 59.6 | 160 | AAE13149 Human stem cell gr |
| 14 | 903 | 59.6 | 160 | ABE13174 Human secreted pro |
| 15 | 903 | 59.6 | 160 | AAW79312 Human protein SEQ |
| 16 | 638 | 42.1 | 229 | AAE13162 Mouse thymospond |
| 17 | 276 | 18.2 | 46 | AAE13155 Human SCR-1 protei |
| 18 | 221 | 14.6 | 42 | AAE13159 Human SCR-1 protei |
| 19 | 213 | 11.6 | 37 | AAE13157 Human SCR-1 protei |
| 20 | 176.5 | 11.6 | 1679 | ABE60498 Drosophila melanog |
| 21 | 176.5 | 11.6 | 1679 | ABE60502 Drosophila melanog |
| 22 | 173.5 | 11.4 | 969 | AAE13152 Patted basic amino |
| 23 | 173.5 | 11.4 | 969 | AAO17358 Human subtilisin I |
| 24 | 162.5 | 10.7 | 132 | AAW08493 Rat serine proteas |
| 25 | 162 | 10.7 | 799 | AAW18073 CD4+ T-lymphocyte |
| 26 | 162 | 10.7 | 799 | AAW18073 T-lymphocyte prote |
| 27 | 162 | 10.7 | 881 | AAW18072 CD4+ T-lymphocyte |
| 28 | 162 | 10.7 | 881 | AAW18072 T-lymphocyte prote |
| 29 | 162 | 10.7 | 915 | AAW52306 Human pro-protein |
| 30 | 162 | 10.7 | 915 | AAW18071 CD4+ T-lymphocyte |
| 31 | 162 | 10.7 | 915 | AAW73932 T-lymphocyte prote |
| 32 | 161 | 10.6 | 288 | AAW34527 hTcP protein fragm |
| 33 | 160.5 | 10.6 | 216 | AAW08492 Rat serine proteas |
| 34 | 160.5 | 10.6 | 479 | AAW16980 Human novel secret |
| 35 | 159 | 10.5 | 193 | AAW08491 Rat serine proteas |
| 36 | 158.5 | 10.5 | 379 | AAW08065 Human EGF-like hom |
| 37 | 158.5 | 10.5 | 379 | AAW13345 Amino acid sequenc |
| 38 | 158.5 | 10.5 | 379 | AAW24397 Human PRO17 prote |
| 39 | 158.5 | 10.5 | 379 | AAW70669 Human PRO17 prote |
| 40 | 158.5 | 10.5 | 379 | AAW44822 Human molecule ass |
| 41 | 158.5 | 10.5 | 379 | AAW12325 Human PRO17 polyp |
| 42 | 158.5 | 10.5 | 379 | AAW00822 Human immune respo |
| 43 | 158.5 | 10.5 | 379 | AAW31185 Amino acid sequenc |
| 44 | 158.5 | 10.5 | 379 | AAW80213 Human PRO17 prote |
| 45 | 158.5 | 10.5 | 379 | AAW53076 Human angiotensin |

ALIGNMENTS

| | | |
|----------|----------|---|
| RESULT 1 | AAE13151 | AAE13151 standard; Protein; 273 AA. |
| ID | AAE13151 | 28-JAN-2002 (first entry) |
| XX | AAE13151 | Human stem cell growth factor-like protein #3. |
| XX | XX | Human, stem cell growth factor-like protein; antiinflammatory; neutrotropic; neuroprotective; vitruine; cytosolic; anticonvulsant; immunostimulant; vasotrophic; vitruine; dermatological; tranquilizer; cerebroprotective; osteoparatic; immunodeficiency syndrome; chronic granulomatous disease; duplicated immunodeficiency syndrome; Wiskott-Aldrich syndrome; AIDS; acquired immune deficiency syndrome; agammaglobulinemia; thalassemia; Gaucher's disease; lysosomal storage disease; mucopolysaccharidosis; adrenal white matter degeneration; anaemia; neurodegenerative disease; Parkinson's disease; Alzheimer's disease; thrombocytopenia; SCID; severe combined immunodeficiency; immune disorder; autoimmune disorder; multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis; autoimmune pulmonary inflammation; cancer; tumour; gene therapy; SCR-1; supporting factor for the proliferation of stem cell. |
| XX | XX | Homo sapiens. |
| OS | XX | MO200177169-A2. |
| PN | XX | 18-OCT-2001. |
| XX | XX | 05-APR-2001; 2001WO-US11208. |
| PF | XX | 05-APR-2000; 2000US-0543774. |
| XX | XX | |
| PR | XX | |

PR 28-JUN-2000; 2000US-215733P.
 PR 09-JAN-2001; 2001US-075756Z.
 PR 05-FEB-2001; 2001US-266614P.
 XX
 PA (HYSE-) HYSEQ INC.
 PA (KIRI) KIRIN BEER KK.
 XX
 PI Tang TY, Labat I, Tillinghast JS, Sinku A, Liu C, Dmanac RT;
 PI Stache-Crain B, Dickson M, Mize NK, Nishikawa M;
 XX
 DR WPI, 2001-657166/75.
 DR N-PSDB; AAD21725.
 XX
 PT Novel stem cell growth factor like polypeptides and polynucleotides for
 PT identifying modulators useful for treating diseases such as Alzheimer's
 PT disease, cancer, rheumatoid arthritis, osteoporosis -
 XX
 PS Claim 28; Page 214-215; 232pp; English.
 XX
 CC The patent discloses novel stem cell growth factor-like proteins and
 CC polynucleotides encoding them. Proteins of the invention are also known
 CC as supporting factor for the proliferation of stem cells (SCR-1). Stem
 CC cell growth factor-like proteins are useful for supporting proliferation
 CC or survival of a stem cell or germ cell which is preferably primordial
 CC germ cell, germ line stem cell, embryonic stem cell, haematopoietic stem
 CC cell, haematopoietic progenitor cell, pluripotent cell or totipotent
 CC cell. The haematopoietic progenitor cell cultured using stem cell
 CC growth factor-like proteins can replace as a graft for the bone marrow
 CC transplantation or cord blood transplantation for treating a variety
 CC of diseases such as immunodeficiency syndrome, chronic granulomatous
 CC disease, duplicated immunodeficiency syndrome, agammaglobulinemia,
 CC Wiskott-Aldrich syndrome, acquired immune deficiency syndrome (AIDS),
 CC thalassemia, haemolytic anaemia due to enzyme defect, congenital
 CC anaemia such as sickle cell anaemia, Gaucher's disease, lysosomal
 CC storage diseases such as mucopolysaccharidosis, adrenal white matter
 CC degeneration, a variety of cancer and tumours. Proteins of the
 CC invention are useful for treating diseases such as Parkinson's
 CC disease, Alzheimer's disease and other neurodegenerative diseases,
 CC thrombocytopaenia, immune deficiencies and disorders such as severe
 CC combined immunodeficiency (SCID) and autoimmune disorders such as
 CC multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis,
 CC and autoimmune pulmonary inflammation. Sequences of the invention are
 CC also useful in gene therapy. The present sequence is stem cell growth
 CC factor-like protein from human.
 CC
 XX
 SO Sequence 273 AA:
 Query Match 100.0%; Score 1516; DB 22; Length 273;
 Best Local Similarity 100.0%; Pred. No. 3e-112;
 Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGHRLRLISLWFIILFMFVIGSONASGRORRMHNVNVOGCGCATCSYNGCLSCRP 60
 DB 1 MGHRLRLISLWFIILFMFVIGSONASGRORRMHNVNVOGCGCATCSYNGCLSCRP 60
 QY 61 RLFFALERIGMKGIVCLSSCPGSGYTRYPDINKTKKADCDCTFNKRNCTKCKSGFP 120
 DB 61 RLFFALERIGMKGIVCLSSCPGSGYTRYPDINKTKKADCDCTFNKRNCTKCKSGFP 120
 QY 121 LHLGKCLDPCPEGLANHTMECVSIHCEVSEPMNPSPCTKKGKTCFPGKGTETRRREI 180
 DB 121 LHLGKCLDPCPEGLANHTMECVSIHCEVSEPMNPSPCTKKGKTCFPGKGTETRRREI 180
 QY 181 IOHPSAKGULCPPTNETKCTVOKKKOKERKKERKKRKKPKGSKAIPDSKSL 240
 DB 181 IOHPSAKGULCPPTNETKCTVOKKKOKERKKERKKRKKPKGSKAIPDSKSL 240
 QY 241 SSKFPEORENKKOQKRRVODKOKSVSVSTVH 273
 DB 241 SSKFPEORENKKOQKRRVODKOKSVSVSTVH 273
 RFSULT 2

AAE13150
 ID AAE13150 standard; Protein: 272 AA.
 XX
 AC AAE13150;
 XX
 DT 28-JAN-2002 (first entry)
 XX
 DE Human stem cell growth factor-like protein #2.
 KW Human; stem cell growth factor-like protein; antiinflammatory; nocotropic;
 KW neuroprotective; vulnerary; cytosolic; anticonvulsant; immunostimulant;
 KW vasotropic; virucide; dermatological; tranquilliser; cerebroprotective;
 KW osteopathic; immunodeficiency syndrome; chronic granulomatous disease;
 KW duplicated immunodeficiency syndrome; Wiskott-Aldrich syndrome; AIDS;
 KW acquired immune deficiency syndrome; agammaglobulinemia; thalassemia;
 KW Gaucher's disease; lysosomal storage disease; mucopolysaccharidosis;
 KW adrenal white matter degeneration; anaemia; neurodegenerative disease;
 KW Parkinson's disease; Alzheimer's disease; thrombocytopaenia; SCID;
 KW severe combined immunodeficiency; immune disorder; autoimmune disorder;
 KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
 KW autoimmune pulmonary inflammation; cancer; tumour; gene therapy;
 KW SCR-1; supporting factor for the proliferation of stem cell.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT Protein /label= Signal_peptide
 FT 22..272
 FT /note= "Human mature stem cell growth
 FT factor-like protein"
 PN W0200177169-A2.
 PD 18-OCT-2001.
 XX
 PF 05-APR-2001; 2001WO-US11206.
 XX
 PR 05-APR-2000; 2000US-0543774.
 PR 28-JUN-2000; 2000US-215733P.
 PR 09-JAN-2001; 2001US-075756Z.
 PR 05-FEB-2001; 2001US-266614P.
 XX
 PA (HYSE-) HYSEQ INC.
 PA (KIRI) KIRIN BEER KK.
 XX
 PI Tang TY, Labat I, Tillinghast JS, Sinku A, Liu C, Dmanac RT;
 PI Stache-Crain B, Dickson M, Mize NK, Nishikawa M;
 XX
 DR WPI, 2001-657166/75.
 DR N-PSDB; AAD21724.
 XX
 PT Novel stem cell growth factor like polypeptides and polynucleotides for
 PT identifying modulators useful for treating diseases such as Alzheimer's
 PT disease, cancer, rheumatoid arthritis, osteoporosis -
 XX
 PS Claim 28; Page 211-212; 232pp; English.
 XX
 CC The patent discloses novel stem cell growth factor-like proteins and
 CC polynucleotides encoding them. Proteins of the invention are also known
 CC as supporting factor for the proliferation of stem cells (SCR-1). Stem
 CC cell growth factor-like proteins are useful for supporting proliferation
 CC or survival of a stem cell or germ cell which is preferably primordial
 CC germ cell, germ line stem cell, embryonic stem cell, haematopoietic stem
 CC cell, haematopoietic progenitor cell, pluripotent cell or totipotent
 CC cell. The haematopoietic progenitor cell cultured using stem cell
 CC growth factor-like proteins can replace as a graft for the bone marrow
 CC transplantation or cord blood transplantation for treating a variety
 CC of diseases such as immunodeficiency syndrome, chronic granulomatous
 CC disease, duplicated immunodeficiency syndrome, agammaglobulinemia,
 CC Wiskott-Aldrich syndrome, acquired immune deficiency syndrome (AIDS),
 CC thalassemia, haemolytic anaemia due to enzyme defect, congenital
 CC anaemia such as sickle cell anaemia, Gaucher's disease, lysosomal

CC storage diseases such as mucopolysaccharidosis, adrenal white matter
 CC degeneration, a variety of cancer and tumours. Proteins of the
 CC invention are useful for treating diseases such as Parkinson's
 CC disease, Alzheimer's disease and other neurodegenerative diseases,
 CC chromocytopenia, immune deficiencies and disorders such as severe
 CC combined immunodeficiency (SCID) and autoimmune disorders such as
 CC multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis,
 CC and autoimmune pulmonary inflammation. Sequences of the invention are
 CC also useful in gene therapy. The present sequence is stem cell growth
 CC factor-like protein from human.

XX
 XX Sequence 272 AA:

Query Match 99.3%; Score 1505; DB 22; Length 272;
 Best Local Similarity 100.0%; Pred. No. 2,3e-111;
 Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HRLISWLFILINFEYIGSONASRGRRORRHPNVSGCCGGCATGSDVNGCLSCPKPL 62
 DB 2 HRLISWLFILINFEYIGSONASRGRRORRHPNVSGCCGGCATGSDVNGCLSCPKPL 61

QY 63 FFALEIRIGMKQIGVCLSSCPGYYGTRYPDINKCTCKACDCTCFNKKFCFKCKSGFYHL 122
 DB 62 FFALEIRIGMKQIGVCLSSCPGYYGTRYPDINKCTCKACDCTCFNKKFCFKCKSGFYHL 121

QY 123 LGKCLDNCEGLNANNHMECVSIHCEVSEWNPSPCTKKGKTCGFRGTERVREIIQ 182
 DB 122 LGKCLDNCEGLNANNHMECVSIHCEVSEWNPSPCTKKGKTCGFRGTERVREIIQ 181

QY 183 HPSAKGNLCPTNETRKTCTVQRKKCKGGRKKRERKPKNGSKESKEALPDSKSLSS 242
 DB 182 HPSAKGNLCPTNETRKTCTVQRKKCKGGRKKRERKPKNGSKESKEALPDSKSLSS 241

QY 243 KEIPEORENKQOQKKRVQDKQKSVSVTVH 273
 DB 242 KEIPEORENKQOQKKRVQDKQKSVSVTVH 272

RESULT 3
 AAE13168
 ID AAE13168 standard; Protein; 272 AA.
 AC AAE13168;
 XX
 XX 28-JAN-2002 (first entry)
 DT
 XX
 XX Human stem cell growth factor-like protein #4.
 DE
 XX
 XX Human; stem cell growth factor-like protein; antiinflammatory; neutropic;
 KW neuroprotective; vulnerary; cytosolic; anticonvulsant; immunostimulant;
 KW vasotrophic; virucide; dermatological; tranquilliser; cerebroprotective;
 KW osteopathic; immunodeficiency syndrome; chronic granulomatous disease;
 KW duplicated immunodeficiency syndrome; Miskott-Aldrich syndrome; AIDS;
 KW acquired immune deficiency syndrome; agammaglobulinemia; thalassemia;
 KW Gaucher's disease; lysosomal storage disease; mucopolysaccharidosis;
 KW adrenial white matter degeneration; anaemia; neurodegenerative disease;
 KW Parkinson's disease; Alzheimer's disease; thrombocytopenia; SCID;
 KW severe combined immunodeficiency; immune disorder; autoimmune disorder;
 KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
 KW autoimmune pulmonary inflammation; cancer; tumour; gene therapy;
 KW SCR-1; supporting factor for the proliferation of stem cell.

OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 1..21
 FT /label= Signal_peptide
 FT Protein 22..272 "Human mature stem cell growth
 FT /note= "Human mature stem cell growth
 XX factor-like protein"
 XX
 XX WO200177169-A2.

PD 18-OCT-2001.
 XX
 PF 05-APR-2001; 2001WO-US11208.
 XX
 PR 05-APR-2000; 2000US-0543774.
 XX
 PR 28-JUN-2000; 2000US-215733P.
 PR 09-JAN-2001; 2001US-0757562.
 PR 05-FEB-2001; 2001US-266614P.
 XX
 PA (HYSE-) HYSEO INC.
 PA (KIRI) KIRIN BEER KK.
 XX
 PI Tang TY, Labat I, Tillinghaet JS, Sinku A, Liu C, Dmanac RT;
 PI Stache-Crain B, Dickson W, Mize NK, Nishikawa M;
 XX
 DR WPI; 2001-657166/75.
 DR N-PSDB; AAD21728.
 PT Novel stem cell growth factor like polypeptides and polynucleotides for
 PT identifying modulators useful for treating diseases such as Alzheimer's
 PT disease, cancer, rheumatoid arthritis, osteoporosis
 XX
 XX Claim 28; Page 226-227; 232pp; English.

CC The patent discloses novel stem cell growth factor-like proteins and
 CC polynucleotides encoding them. Proteins of the invention are also known
 CC as supporting factor for the proliferation of stem cells (SCR-1). Stem
 CC cell growth factor-like proteins are useful for supporting proliferation
 CC or survival of a stem cell or germ cell which is preferably, primordial
 CC germ cell, germ line stem cell, embryonic stem cell, haematopoietic stem
 CC cell, haematopoietic progenitor cell, pluripotent cell or totipotent
 CC cell. The haematopoietic progenitor cell cultured using stem cell
 CC growth factor-like proteins can replace as a graft for the bone marrow
 CC transplantation or cord blood transplantation for treating a variety
 CC of diseases such as immunodeficiency syndrome, chronic granulomatous
 CC disease, duplicated immunodeficiency syndrome, agammaglobulinemia,
 CC Miskott-Aldrich syndrome, acquired immune deficiency syndrome (AIDS),
 CC thalassemia, haemolytic anaemia due to enzyme defect, congenital
 CC anaemia such as sickle cell anaemia, Gaucher's disease, lysosomal
 CC storage diseases such as mucopolysaccharidosis, adrenal white matter
 CC degeneration, a variety of cancer and tumours. Proteins of the
 CC invention are useful for treating diseases such as Parkinson's
 CC disease, Alzheimer's disease and other neurodegenerative diseases,
 CC chromocytopenia, immune deficiencies and disorders such as severe
 CC combined immunodeficiency (SCID) and autoimmune disorders such as
 CC multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis,
 CC and autoimmune pulmonary inflammation. Sequences of the invention are
 CC also useful in gene therapy. The present sequence is stem cell growth
 CC factor-like protein from human.

XX
 XX Sequence 272 AA:
 SO

Query Match 99.3%; Score 1505; DB 22; Length 272;
 Best Local Similarity 100.0%; Pred. No. 2,3e-111;
 Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HRLISWLFILINFEYIGSONASRGRRORRHPNVSGCCGGCATGSDVNGCLSCPKPL 62
 DB 2 HRLISWLFILINFEYIGSONASRGRRORRHPNVSGCCGGCATGSDVNGCLSCPKPL 61

QY 63 FFALEIRIGMKQIGVCLSSCPGYYGTRYPDINKCTCKACDCTCFNKKFCFKCKSGFYHL 122
 DB 62 FFALEIRIGMKQIGVCLSSCPGYYGTRYPDINKCTCKACDCTCFNKKFCFKCKSGFYHL 121

QY 123 LGKCLDNCEGLNANNHMECVSIHCEVSEWNPSPCTKKGKTCGFRGTERVREIIQ 182
 DB 122 LGKCLDNCEGLNANNHMECVSIHCEVSEWNPSPCTKKGKTCGFRGTERVREIIQ 181

QY 183 HPSAKGNLCPTNETRKTCTVQRKKCKGGRKKRERKPKNGSKESKEALPDSKSLSS 242
 DB 182 HPSAKGNLCPTNETRKTCTVQRKKCKGGRKKRERKPKNGSKESKEALPDSKSLSS 241

QY 243 KEIPEORENKQOQKKRVQDKQKSVSVTVH 273


```

|||||
Db 2 HRLRLSWLFIILNFMETIGSONASRGRRMRHNPVSGCCGGCATGSDYNGLSCKPRL 61
Qy 63 PALERIRKMGKQIGVCLSSCPGSGYGRYPDINKCTKACADDTCTFNKFTCKSGPYLH 122
Db 63 PALERIRKMGKQIGVCLSSCPGSGYGRYPDINKCTKACADDTCTFNKFTCKSGPYLH 121
Qy 123 LGKCLDNCPEGLLEANNHTMECVSIHVCEVSENNPMSPTCKGKTCGKGTETRVREIIQ 182
Db 122 LGKCLDNCPEGLLEANNHTMECVSIHVCEVSENNPMSPTCKGKTCGKGTETRVREIIQ 181
Qy 183 HPSAKGNLCPTNTRKCTVORRKCCKGGRKKGERKKRKNPKNGESKEAIPDSKLESS 242
Db 182 HPSAKGNLCPTNTRKCTVORRKCCKGGRKKGERKKRKNPKNGESKEAIPDSKLESS 241
Qy 243 KEIPEORENKQOQKKRKVDKOKSVSVSTVH 273
Db 242 KEIPEORENKQOQKKRKVDKOKSVSVSTVH 272

RESULT 6
AAE13163
ID AAE13163 standard; Protein; 265 AA.
AC AAE13163;
XX
XX
DT 28-JAN-2002 (first entry)
DE Human secreted protein from clone DA228_6.
XX
XX
Human; stem cell growth factor-like protein; antiinflammatory; neutrophic;
neuroprotective; vulnary; cytosolic; anticonvulsant; immunostimulant;
vasotrophic; vitruide; dermatological; tranquiliser; cerebroprotective;
osteopathic; immunodeficiency syndrome; chronic granulomatous disease;
duplicated immunodeficiency syndrome; Miskot-Aldrich syndrome; AIDS;
acquired immune deficiency syndrome; agammaglobulinemia; thalassemia;
Gaucher's disease; lysosomal storage disease; mucopolysaccharidosis;
adrenal white matter degeneration; anaemia; neurodegenerative disease;
Parkinson's disease; Alzheimer's disease; thrombocytopenia; SCID;
severe combined immunodeficiency; immune disorder; autoimmune disease;
multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
autoimmune pulmonary inflammation; cancer; tumour; gene therapy; SCR-1;
supporting factor for the proliferation of stem cell; secreted protein.
XX
OS Homo sapiens.
XX
XX
PN WO200177169-A2.
PD 18-OCT-2001.
XX
XX
PF 05-APR-2001; 2001WO-US11208.
XX
XX
PR 05-APR-2000; 2000US-0543774.
PR 28-JUN-2000; 2000US-215733P.
PR 09-JAN-2001; 2001US-0757562.
PR 05-FEB-2001; 2001US-266614P.
XX
XX
PA (HYSE-) HYSEQ INC.
PA (KIRI) KIRIN BEER KK.
XX
XX
PI Tang TY, Labat I, Tillinghast JS, Sinku A, Liu C, Drmanac RT;
PI Stache-Crain B, Dickson M, Mize NK, Nishikawa M;
XX
XX
DR WPI: 2001-657166/75.
XX
XX
PT Novel stem cell growth factor like polypeptides and polynucleotides for
PT identifying modulators useful for treating diseases such as Alzheimer's
PT disease, cancer, rheumatoid arthritis, osteoporosis
XX
XX
PS Disclosure; Fig 3; 232pp. English.
XX
XX
CC The patent discloses novel stem cell growth factor-like proteins and
CC polynucleotides encoding them. Proteins of the invention are also known

```

```

CC as supporting factor for the proliferation of stem cells (SCR-1). Stem
CC cell growth factor-like proteins are useful for supporting proliferation
CC or survival of a stem cell or germ cell which is preferably primordial
CC germ cell, germ line stem cell, embryonic stem cell, haematopoietic stem
CC cell, haematopoietic progenitor cell, pluripotent cell or totipotent
CC cell. The haematopoietic progenitor cell cultured using stem cell
CC growth factor-like proteins can replace as a graft for the bone marrow
CC transplantation or cord blood transplantation for treating a variety
CC of diseases such as immunodeficiency syndrome, chronic granulomatous
CC disease, duplicated immunodeficiency syndrome, agammaglobulinemia,
CC Miskot-Aldrich syndrome, acquired immune deficiency syndrome (AIDS),
CC thalassemia, haemolytic anaemia due to enzyme defect, congenital
CC anaemia such as sickle cell anaemia, Gaucher's disease, lysosomal
CC storage diseases such as mucopolysaccharidosis, adrenal white matter
CC degeneration, a variety of cancer and tumours. Proteins of the
CC invention are useful for treating diseases such as Parkinson's
CC disease, Alzheimer's disease and other neurodegenerative diseases,
CC thrombocytopenia, immune deficiencies and disorders such as severe
CC combined immunodeficiency (SCID) and autoimmune disorders such as
CC multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis,
CC and autoimmune pulmonary inflammation. Sequences of the invention are
CC also useful in gene therapy. The present sequence is human secreted
CC protein from clone DA228_6.
XX
SQ Sequence 265 AA:
Query Match 97.1%; Score 1472; DB 22; Length 265;
Best Local Similarity 100.0%; Pred. No. 9e-109;
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 HRLRLSWLFIILNFMETIGSONASRGRRMRHNPVSGCCGGCATGSDYNGLSCKPRL 62
Db 2 HRLRLSWLFIILNFMETIGSONASRGRRMRHNPVSGCCGGCATGSDYNGLSCKPRL 61
Qy 63 PALERIRKMGKQIGVCLSSCPGSGYGRYPDINKCTKACADDTCTFNKFTCKSGPYLH 122
Db 62 PALERIRKMGKQIGVCLSSCPGSGYGRYPDINKCTKACADDTCTFNKFTCKSGPYLH 121
Qy 123 LGKCLDNCPEGLLEANNHTMECVSIHVCEVSENNPMSPTCKGKTCGKGTETRVREIIQ 182
Db 122 LGKCLDNCPEGLLEANNHTMECVSIHVCEVSENNPMSPTCKGKTCGKGTETRVREIIQ 181
Qy 183 HPSAKGNLCPTNTRKCTVORRKCCKGGRKKGERKKRKNPKNGESKEAIPDSKLESS 242
Db 182 HPSAKGNLCPTNTRKCTVORRKCCKGGRKKGERKKRKNPKNGESKEAIPDSKLESS 241
Qy 243 KEIPEORENKQOQKKRKVDKOKSVSVSTVH 273
Db 242 KEIPEORENKQOQKKRKVDKOKSVSVSTVH 265

RESULT 7
AAW85607
ID AAW85607 standard; Protein; 292 AA.
XX
XX
AC AAW85607;
XX
XX
DT 02-MAR-1999 (first entry)
DE Secreted protein clone da228_6.
XX
XX
XX
XX
PI Clone; secreted protein; protein factor; cytokine; lymphokine;
PI interferon; colony stimulating factor; CSF; interleukin; cloning;
XX
XX
KM tumour invasion; tumour suppression; immune boosting.
XX
XX
OS Homo sapiens.
XX
XX
PN WO9849302-A1.
XX
XX
PD 05-NOV-1998.
XX
XX
PF 24-APR-1998; 98WO-US08336.
XX
XX

```

PR 23-APR-1998; 98US-0065125.
PR 25-APR-1997; 97US-0845296.
XX
PA (GENY) GENETICS INST INC.
XX
PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
PI Racie LA, Spaulding V, Treacy M;
DR WPI; 1999-024059/02.
DR N-PSDB; AAV83133.
XX
XX New polynucleotides encoding secreted human proteins - are derived
PT from human foetal brain, adult brain, adult blood or placenta cDNA
PT libraries, useful, e.g. as potential immunomodulators
XX
PS Claim 8; Page 63-64; 104pp; English.
XX
XX The nucleotide sequence (NS) of the full-length protein-coding
CC sequence of clones c1254 (AAV83132), da2286 (AAV83133), du4105
CC (AAV83134), eh801 (AAV83135), e73691 (AAV83136), fh1235 (AAV83137), fm601
CC (AAV83138) or fr4732 (AAV83139), (all clones are deposited as ATCC 98415)
CC and the proteins they encode are predicted to have biological activities
CC which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals for example,
CC tumour suppression/invasion activity, immune system boosting
CC activity. The polynucleotides are also believed to be useful for
CC gene therapy.
XX
SQ Sequence 292 AA;
XX
Query Match 97.1%; Score 1472; DB 20; Length 292;
Best Local Similarity 100.0%; Pred. No. 1e-108;
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HRLRISWLFILNMEYIGSONASRGRRORRHPNVSGCCGCGCATGSDYNGCLSCRPRL 62
DB 2 HRLRISWLFILNMEYIGSONASRGRRORRHPNVSGCCGCGCATGSDYNGCLSCRPRL 61
QY 63 FFALRIGMKOIGVCLSSCPGSGYGTGTRYPDINKCTKACADCTCFNNKFCCKKSGFYHL 122
DB 62 FFALRIGMKOIGVCLSSCPGSGYGTGTRYPDINKCTKACADCTCFNNKFCCKKSGFYHL 121
QY 123 LKCLDNCPEGLEANNHTMECVSIHCEVSENNPWSPTCKKGTGCRGCTERRVREIIIO 182
DB 122 LKCLDNCPEGLEANNHTMECVSIHCEVSENNPWSPTCKKGTGCRGCTERRVREIIIO 181
QY 183 HPSAKGNLCPTNETRKTCTVORKKCKGGERGKKGRERKPKPNKGESKEAIPDSKLESS 242
DB 182 HPSAKGNLCPTNETRKTCTVORKKCKGGERGKKGRERKPKPNKGESKEAIPDSKLESS 241
QY 243 KEIPEORENKQOKKRYVODKXS 266
DB 242 KEIPEORENKQOKKRYVODKXS 265

RESULT 8
AAE13170
ID AAE13170 standard; Protein; 292 AA.
XX
XX
AC AAE13170;
XX
XX
DT 28-JAN-2002 (first entry)
XX
XX Human SCR-1 related protein.
DE
XX Human; stem cell growth factor-like protein; antiinflammatory; nocotropic;
KM neuroprotective; vulnerary; cyostatic; anticonvulsant; immunostimulant;
KM vasotrophic; virucide; dermatological; tranquilliser; cerebroprotective;
KM osteoplastic; immunodeficiency syndrome; chronic granulomatous disease;
KM duplicated immunodeficiency syndrome; Miskot-Aldrich syndrome; AIDS;
KM acquired immune deficiency syndrome; agammaglobulinemia; thalassemia;
KM Gaucher's disease; lysosomal storage disease; mucopolysaccharidosis;
KM adrenal white matter degeneration; anaemia; neurodegenerative disease;

KM Parkinson's disease; Alzheimer's disease; thrombocytopaenia; SCID;
KM severe combined immunodeficiency; immune disorder; autoimmune disorder;
KM multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
KM autoimmune pulmonary inflammation; cancer; tumour; gene therapy;
KM SCR-1; supporting factor for the proliferation of stem cell.
XX
XX unidentified.
OS
XX WO200177169-A2.
XX
XX 18-OCT-2001.
XX
XX 05-APR-2001; 2001WO-US11208.
XX
XX 05-APR-2000; 2000US-0543774.
XX 28-JUN-2000; 2000US-215733P.
PR 09-JAN-2001; 2001US-0757562.
PR 05-FEB-2001; 2001US-266614P.
XX
XX (HYSE-) HYSEQ INC.
XX (KIRI) KIRIN BEER KK.
XX
XX Tang TY, Labat I, Tillinghast JS, Sinku A, Liu C, Drmanac RT;
PI Stache-Crain B, Dickson M, Mize NK, Nishikawa M;
DR WPI; 2001-657166/75.
DR N-PSDB; AAD21740.
XX
XX Novel stem cell growth factor like polypeptides and polynucleotides for
PT identifying modulators useful for treating diseases such as Alzheimer's
PT disease, cancer, rheumatoid arthritis, osteoporosis -
XX
XX Claim 27; Page 231-232; 232pp; English.
XX
XX The patent discloses novel stem cell growth factor-like proteins and
CC polynucleotides encoding them. Proteins of the invention are also known
CC as supporting factor for the proliferation of stem cells (SCR-1). Stem
CC cell growth factor-like proteins are useful for supporting proliferation
CC or survival of a stem cell or germ cell which is preferably primordial
CC germ cell, germ line stem cell, embryonic stem cell, haematopoietic stem
CC cell, haematopoietic progenitor cell, pluripotent cell or totipotent
CC cell. The haematopoietic progenitor cell cultured using stem cell
CC growth factor-like proteins can replace as a graft for the bone marrow
CC transplantation or cord blood transplantation for treating a variety
CC of diseases such as immunodeficiency syndrome, chronic granulomatous
CC disease, duplicated immunodeficiency syndrome, acquired immune deficiency
CC disease, Miskot-Aldrich syndrome, acquired immune deficiency syndrome (AIDS),
CC thalassemia, haemolytic anaemia due to enzyme defect, congenital
CC anaemia such as sickle cell anaemia, Gaucher's disease, lysosomal
CC storage diseases such as mucopolysaccharidosis, adrenal white matter
CC degeneration, a variety of cancer and tumours. Proteins of the
CC invention are useful for treating diseases such as Parkinson's
CC disease, Alzheimer's disease and other neurodegenerative diseases,
CC thrombocytopaenia, immune deficiencies and disorders such as severe
CC combined immunodeficiency (SCID) and autoimmune disorders such as
CC multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis,
CC and autoimmune pulmonary inflammation. Sequences of the invention are
CC also useful in gene therapy. The present sequence is human SCR-1
CC related protein.
XX
XX
SQ Sequence 292 AA;
XX
Query Match 97.1%; Score 1472; DB 22; Length 292;
Best Local Similarity 100.0%; Pred. No. 1e-108;
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HRLRISWLFILNMEYIGSONASRGRRORRHPNVSGCCGCGCATGSDYNGCLSCRPRL 62
DB 2 HRLRISWLFILNMEYIGSONASRGRRORRHPNVSGCCGCGCATGSDYNGCLSCRPRL 61
QY 63 FFALRIGMKOIGVCLSSCPGSGYGTGTRYPDINKCTKACADCTCFNNKFCCKKSGFYHL 122
DB 62 FFALRIGMKOIGVCLSSCPGSGYGTGTRYPDINKCTKACADCTCFNNKFCCKKSGFYHL 121

QY 123 LGKCLDNCPEGLANNHTMECVSIHCEVSENNPWSPTCKGKTGCFRTETRVREIIQ 182
DB 122 LGKCLDNCPEGLANNHTMECVSIHCEVSENNPWSPTCKGKTGCFRTETRVREIIQ 181
QY 183 HPSAKGNLCPTNETRKTCTVORRKCCKGGRGKGRERKPKNGSKESKEAIPDSKSLSS 242
DB 182 HPSAKGNLCPTNETRKTCTVORRKCCKGGRGKGRERKPKNGSKESKEAIPDSKSLSS 241
QY 243 KEIPEORENKQOQKKRKYODKOKS 266
DB 242 KEIPEORENKQOQKKRKYODKOKS 265

RESULT 9
ABP61846
ID ABP61846 standard; Protein; 292 AA.
XX
AC ABP61846;
XX
DT 04-OCT-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 200.
XX
KW Human; cytosolic; antirheumatic; antiarthritic; vulnery; analgesic;
KW antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;
KW neuroprotective; nocotropic; osteopathic; haemostatic; vasotropic;
KW antitumor; fungicide; antidiabetic; antiallergic;
KW immunostimulant; antiparasitic; secreted protein; transmembrane protein;
KW cytokine; cell proliferation; cell differentiation; autoimmune disease;
KW stem cell; growth factor; nervous system disease; neuropathy;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW osteoporosis; severe combined immunodeficiency; SCID; infection;
KW multiple sclerosis; rheumatoid arthritis; gene therapy.
XX
OS Homo sapiens.
XX
PN US2002065394-A1.
XX
PD 30-MAY-2002.
XX
PF 22-DEC-2000; 2000US-0745763.
XX
PR 18-MAR-1998; 98US-0040963.
XX
PA (JACO/) JACOBS K.
PA (MCCO/) MCCOY J.M.
PA (LAVA/) LAVALLIE E.R.
PA (COLL/) COLLINS-RACIE L.A.
PA (EVAN/) EVANS C.
PA (MERB/) MERBERG D.
PA (TREA/) TREACY M.
PA (SPAU/) SPAULDING V.
PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI Metzberg D, Treacy M, Spaulding V;
XX
DR WPI; 2002-582343/62.
XX
DR N-PSDB; ABO92060.
XX
PT Novel secreted or transmembrane protein and polynucleotide encoding the
PT protein, useful for diagnosis and treatment of neurological disorders,
PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis
PT
XX
PS Claim 207, Page 203-204; 284pp; English.
XX
XX The invention relates to human secreted or transmembrane protein (I),
CC their fragments and is encoded by specific complementary deoxyribonucleic
CC acid (CDNA) inserts (II), where the protein is substantially free from
CC other mammalian proteins, (I) are useful for preventing, treating or
CC ameliorating a medical condition, especially immunological treatment or
CC prevention of tumours. (I) exhibits activity relating to angiogenesis.

CC cytokine, cell proliferation, cell differentiation, antiinflammatory,
CC stem cell growth factor activity and activin or inhibin-related
CC activities. (I) can be used to manipulate stem cells in culture to give
CC rise to neuroepithelial cells that can be used to augment or replace
CC cells damaged by illness, autoimmune disease, accidental damage or
CC genetic disorders. (I) induces the proliferation of neural cells and
CC regeneration of nerve and brain tissue and is useful for the treatment of
CC central and peripheral nervous system diseases and neuropathies, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis. (II) is involved in chemotactic or chemokinetic
CC activity, regulation of hematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,
CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or
CC periodontal disease. (I) is also useful for gut protection or
CC regeneration and treatment of lung or liver fibrosis, reperfusion injury
CC in various tissues, various immune deficiencies and disorders including
CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,
CC such as asthma or other respiratory problems. (II) is useful to express
CC recombinant protein, as markers for tissues in which the corresponding
CC protein is preferentially expressed and in gene therapy. The present
CC sequence is that of a polypeptide of the invention.
XX
SO Sequence 292 AA.
XX
Query Match 97.1%; Score 1472; DB 23; Length 292;
Best Local Similarity 100.0%; Pred. No. 1e-108;
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 HRLISLWLTIIINFEVYIGSONASRGRRORRMHPNWSQCGCGCATCSDYNGCLSCPKRL 62
DB 2 HRLISLWLTIIINFEVYIGSONASRGRRORRMHPNWSQCGCGCATCSDYNGCLSCPKRL 61
QY 63 FFLERIGMKQIGVCLSCPSGYGTRYPDIKCTKCAADDTCENKPFCKKSGFYH 122
DB 62 FFLERIGMKQIGVCLSCPSGYGTRYPDIKCTKCAADDTCENKPFCKKSGFYH 121
QY 123 LGKCLDNCPEGLANNHTMECVSIHCEVSENNPWSPTCKGKTGCFRTETRVREIIQ 182
DB 122 LGKCLDNCPEGLANNHTMECVSIHCEVSENNPWSPTCKGKTGCFRTETRVREIIQ 181
QY 183 HPSAKGNLCPTNETRKTCTVORRKCCKGGRGKGRERKPKNGSKESKEAIPDSKSLSS 242
DB 182 HPSAKGNLCPTNETRKTCTVORRKCCKGGRGKGRERKPKNGSKESKEAIPDSKSLSS 241
QY 243 KEIPEORENKQOQKKRKYODKOKS 266
DB 242 KEIPEORENKQOQKKRKYODKOKS 265

RESULT 10
AAB93875
ID AAB93875 standard; Protein; 292 AA.
XX
AC AAB93875;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:13761.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX

PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J,
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs.
 XX
 PS Claim 8; SEQ ID 13781; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises at least 15 nucleotides and the combination
 CC of the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SO Sequence 292 AA;
 Query Match 95.8%; Score 1452; DB 22; Length 292;
 Best Local Similarity 99.2%; Pred. No. 3.8e-107;
 Matches 262; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 3 HRLISMLFIILNMEYIGSONASRGRORRMHPVNSGCGGATGSDYNGLSCKPRL 62
 DB 2 HRLISMLFIILNMEYIGSONASRGRORRMHPVNSGCGGATGSDYNGLSCKPRL 61
 QY 63 PALERIGMKOIGVCLSSGSGYGYRRYPDINKCTKCAADCTCPNKPACTCKSGFYAH 122
 DB 62 PALERIGMKOIGVCLSSGSGYGYRRYPDINKCTKCAADCTCPNKPACTCKSGFYAH 121
 QY 123 LCKCLDNCPEGLEANNHMECVSIHVCESEWNPSPCTKKGKTCGFRGTETRVREIIQ 182
 DB 122 LCKCLDNCPEGLEANNHMECVSIHVCESEWNPSPCTKKGKTCGFRGTETRVREIIQ 181
 QY 183 HSAAGNLCPTNTRKCTVQRKKCKGGRGKGRKKKPPNGSEKXKIPDSKLESS 242
 DB 182 HPSAGNLCPTNTRKCTVQRKKCKGGRGKGRKKKPPNGSEKXKIPDSKLESS 241
 QY 243 KEIPQRENKQOQKKRKYVDQKXS 266
 DB 242 KEIPQRENKQOQKKRKYVDQKXS 265
 RESULT 11
 AAE13153
 IF AAE13153 standard; Protein; 251 AA.

XX
 AC AAE13153;
 XX
 DT 28-JAN-2002 (first entry)
 XX
 DE Human mature stem cell growth factor-like protein.
 XX
 KW Human; stem cell growth factor-like protein; antiinflammatory; neutrophic;
 KW neuroprotective; vulnerary; cytosolic; anticonvulsant; immunostimulant;
 KW vasotrophic; virutide; dermatological; tranquilliser; cerebroprotective;
 KW osteopathic; immunodeficiency syndrome; chronic granulomatous disease;
 KW duplicated immunodeficiency syndrome; Wiskott-Aldrich syndrome; AIDS;
 KW acquired immune deficiency syndrome; agammaglobulinaemia; thalassemia;
 KW Gaucher's disease; lysosomal storage disease; mucopolysaccharidosis;
 KW adental white matter degeneration; anaemia; neurodegenerative disease;
 KW Parkinson's disease; Alzheimer's disease; thrombocytopaenia; SCID;
 KW severe combined immunodeficiency; immune disorder; autoimmune disorder;
 KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
 KW autoimmune pulmonary inflammation; cancer; tumour; gene therapy;
 KW SCR-1; supporting factor for the proliferation of stem cell.
 XX
 OS Homo sapiens.
 XX
 PN WO200177169-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 05-APR-2001; 2001WO-US11208.
 XX
 PR 05-APR-2000; 2000US-0543774.
 PR 28-JUN-2000; 2000US-215733P.
 PR 09-JAN-2001; 2001US-0752562.
 PR 05-FEB-2001; 2001US-266614P.
 XX
 PA (HSE-) HSEQ INC.
 PA (KIRI) KIRIN BEER KK.
 PI Tang TY, Labat I, Tillinghast JS, Sinku A, Liu C, Drmanac RT;
 PI Stache-Crain B, Dickson M, Mize NK, Nishikawa M;
 XX
 DR WPI; 2001-657166/75.
 XX
 PT Novel stem cell growth factor like polypeptides and polynucleotides for
 PT identifying modulators useful for treating diseases such as Alzheimer's
 PT disease, cancer, rheumatoid arthritis, osteoporosis
 XX
 PS Claim 28; Page 216-217; 232pp; English.
 CC The patent discloses novel stem cell growth factor-like proteins and
 CC polynucleotides encoding them. Proteins of the invention are also known
 CC as supporting factor for the proliferation of stem cells (SCR-1). Stem
 CC cell growth factor-like proteins are useful for supporting proliferation
 CC or survival of a stem cell or germ cell which is preferably primordial
 CC germ cell, germ line stem cell, embryonic stem cell, haematopoietic stem
 CC cell, haematopoietic progenitor cell, pluripotent cell or totipotent
 CC cell. The haematopoietic progenitor cell cultured using stem cell
 CC growth factor-like proteins can replace as a graft for the bone marrow
 CC transplantation or cord blood transplantation for treating a variety
 CC of diseases such as immunodeficiency syndrome, chronic granulomatous
 CC disease, duplicated immunodeficiency syndrome, agammaglobulinaemia,
 CC Wiskott-Aldrich syndrome, acquired immune deficiency syndrome (AIDS),
 CC thalassemia, haemolytic anaemia due to enzyme defect, congenital
 CC anaemia such as sickle cell anaemia, Gaucher's disease, lysosomal
 CC storage diseases such as mucopolysaccharidosis, adental white matter
 CC degeneration, a variety of cancer and tumours. Proteins of the
 CC invention are useful for treating diseases such as Parkinson's
 CC disease, Alzheimer's disease and other neurodegenerative diseases,
 CC thrombocytopaenia, immune deficiencies and disorders such as severe
 CC combined immunodeficiency (SCID) and autoimmune disorders such as
 CC multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis,
 CC and autoimmune pulmonary inflammation. Sequences of the invention are
 CC also useful in gene therapy. The present sequence is the mature protein
 CC of human stem cell growth factor-like protein.

| XX | Sequence | 251 AA; | 92.3%; | Score 1400; | DB 22; | Length 251; |
|----|---|--|---------------------|-------------|---------|-------------|
| XX | Query Match | | | | | |
| XX | Best Local Similarity | 100.0%; | Fred. No. 4,2e-103; | | | |
| XX | Matches 251; Conservative | 0; | Mismatches 0; | Indels 0; | Gaps 0; | |
| Qy | 1 | QNASRGRORRMRHPVWSQSGCGGATCSQDYNGLSCSKRPLFFALERIGMKOIGVCLSSCP | 82 | | | |
| Db | 1 | QNASRGRORRMRHPVWSQSGCGGATCSQDYNGLSCSKRPLFFALERIGMKOIGVCLSSCP | 60 | | | |
| Qy | 83 | SGYGGRYPDINKCTCKKADODTCFPHNCTCKCKSGFYLHGLGCLDNCEGLSEANNHME | 142 | | | |
| Db | 61 | SGYGGRYPDINKCTCKKADODTCFPHNCTCKCKSGFYLHGLGCLDNCEGLSEANNHME | 120 | | | |
| Qy | 143 | CSVIVHCEVSENNPMSPTCKKGTCCFPGKGTETTRVREIIGHPSAKNLCPPNETRKCTV | 202 | | | |
| Db | 121 | CSVIVHCEVSENNPMSPTCKKGTCCFPGKGTETTRVREIIGHPSAKNLCPPNETRKCTV | 180 | | | |
| Qy | 203 | QRKKCKGGRGKGRERKRRKPKNGSKSEALPDSKLSSESKXEPEDRENNKQOKKRYOD | 262 | | | |
| Db | 181 | QRKKCKGGRGKGRERKRRKPKNGSKSEALPDSKLSSESKXEPEDRENNKQOKKRYOD | 240 | | | |
| Qy | 263 | KOKSVSVSTVH | 273 | | | |
| Db | 241 | KOKSVSVSTVH | 251 | | | |
| XX | RESULT 12 | | | | | |
| XX | AAEL13167 | | | | | |
| XX | AAEL13167 | standard; Protein; 279 AA. | | | | |
| XX | AAEL13167; | | | | | |
| XX | 28-JAN-2002 | (first entry) | | | | |
| DE | Mouse stem cell growth factor-like protein. | | | | | |
| KW | Mouse; stem cell growth factor-like protein; antiinflammatory; neotrophic; neuroprotective; vulnary; cytostatic; anticonvulsant; immunostimulant; vasotropic; vituric; dermatological; tranquilliser; cerebroprotective; osteopathic; immunodeficiency syndrome; chronic granulomatous disease; duplicated immunodeficiency syndrome; Wiskott-Aldrich syndrome; AIDS; acquired immune deficiency syndrome; agammaglobulinaemia; thalassemia; Gaucher's disease; lysosomal storage disease; mucopolysaccharidoses; adrenal white matter degeneration; anaemia; neurodegenerative disease; Parkinson's disease; Alzheimer's disease; chromocytopenia; SCID; severe combined immunodeficiency; immune disorder; autoimmune disorder; multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis; autoimmune pulmonary inflammation; cancer; tumour; gene therapy; SCR-1; supporting factor for the proliferation of stem cell. | | | | | |
| XX | Mus musculus. | | | | | |
| XX | Key | Location/Qualifiers | | | | |
| XX | Peptide | 1..21 | | | | |
| FT | Protein | /label= Signal peptide | | | | |
| FT | | 22..279 | | | | |
| FT | | /note= "Mouse mature stem cell growth factor-like protein" | | | | |
| XX | WO200177169-A2. | | | | | |
| XX | 18-OCT-2001. | | | | | |
| XX | 05-APR-2001; | 2001WO-US11208. | | | | |
| XX | 05-APR-2000; | 2000US-0543774. | | | | |
| XX | 28-JUN-2000; | 2000US-215731P. | | | | |
| XX | 09-JAN-2001; | 2001US-0757562. | | | | |
| XX | 05-FEB-2001; | 2001US-266614P. | | | | |
| XX | (HYSE-) HYSEQ INC. | | | | | |

(KIRI) KIRIN BEER KK.

Teng TY, Labat I, Tillinghaer JS, Slink A, Liu C, Dimaac RT;
Stache-Cran B, .Dickson M, Mize NK, Nishikawa M;
MPI; 2001-657166/75.
DR N.-PSDB; AAD2127.

Novel stem cell growth factor like polypeptides and polynucleotides for identifying modulators useful for treating diseases such as Alzheimer's disease, cancer, rheumatoid arthritis, osteoporosis

Claim 28; Page 223-224; 232pp; English.

The patent discloses novel stem cell growth factor-like proteins and polynucleotides encoding them. Proteins of the invention are also known as supporting factor for the proliferation of stem cells (SCR-1). Stem cell growth factor-like proteins are useful for supporting proliferation or survival of a stem cell or germ cell which is preferably primordial germ cell, germ line stem cell, embryonic stem cell, haematopoietic stem cell, haematopoietic progenitor cell, pluripotent cell or totipotent cell. The haematopoietic progenitor cell cultured using stem cell growth factor-like proteins can replace as a graft for the bone marrow transplantation or cord blood transplantation for treating a variety of diseases such as immunodeficiency syndrome, chronic granulomatous disease, duplicated immunodeficiency syndrome, agammaglobulinemia, Wiskott-Aldrich syndrome, acquired immune deficiency syndrome (AIDS), thalassemia, haemolytic anaemia due to enzyme defect, congenital anaemia such as sickle cell anaemia, Gaucher's disease, lysosomal storage diseases such as mucopolysaccharidosis, adrenal white matter degeneration, a variety of cancer and tumours. Proteins of the invention are useful for treating diseases such as Parkinson's disease, Alzheimer's disease and other neurodegenerative diseases, thrombocytopoenia, immune deficiencies and disorders such as severe combined immunodeficiency (SCID) and autoimmune disorders such as multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, CC and autoimmune pulmonary inflammation. Sequences of the invention are also useful in gene therapy. The present sequence is stem cell growth factor-like protein from mouse.

Sequence 279 AA;

Query Match 86.4%; Score 1310.5; DB 22; Length 279;
Best Local Similarity 87.1%; Pred. No. 5,8e-96;
Matches 242; Conservative 11; Mismatches 18; Indels 7; Gaps 4

3 HURLISLPTILNMEYIGSNASRRORMHPNVSGCGGACATCDNYGNCISCKPRL 62
Db 2 HURLISCFITILNMEYIGSNASRRORMHPNVSGCGGACATCDNYGNCISCKPRL 61

63 PALERIIMKOIGVCLSSCPSPGYGTRPRDINKCTCKAKDCDCEFNKKFCRCKSGFYLL 122
Db 62 PVLIERIKMGKDIGVCLSSCPSPGYGRPRDINKCTCKAKDCDCEFNKKFCRCKSGFYLL 121

123 LQKCIDNPPELEANNHTECVSIHVCSFPMNWSQCTKKGTCGFMRGETVRFLIQ 182
Db 122 LQKCIDSCPBELEANNHTECVSIHVCSFPMNWSQCMKGGTGTGFRGTETRYRDILQ 181

183 HFSA--KGNLCPTNETRKCTVORRKKCKGSRGKGERKKPKNGESKE--AIPOSKS 238
Db 182 HFSAGKGNLCPTNETRKCTVORRKKCKGSRGKGERKKPKNKKERKETSTSSSDSKG 241

239 LESEKEIPRENNK--QQOKKKRVODK-QKSVSVSYTH 273
Db 242 LESEIETPDOENKEROQQOKRARARDKQKSVSVSYTH 279

RESULT 13
AAEI3149
ID AAEI3149 standard; Protein; 160 AA.
XX AAEI3149;
XX

DT 28-JAN-2002 (first entry)

XX Human stem cell growth factor-like protein #1.

XX

KM Human; stem cell growth factor-like protein; antiinflammatory; neutrophic;
 KM neuroprotective; vulnery; cytosolic; anticonvulsant; immunostimulant;
 KM vasotropic; virucide; dermatological; tranquilizer; cerebroprotective;
 KM osteopathic; immunodeficiency syndrome; chronic granulomatous disease;
 KM duplicated immunodeficiency syndrome; Miskott-Aldrich syndrome; AIDS;
 KM acquired immune deficiency syndrome; agammaglobulinemia; thalassemia;
 KM Gaucher's disease; lysosomal storage disease; mucopolysaccharidosis;
 KM adrenal white matter degeneration; anaemia; neurodegenerative disease;
 KM Parkinson's disease; Alzheimer's disease; thrombocytopaenia; SCID;
 KM severe combined immunodeficiency; immune disorder; autoimmune disorder;
 KM multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
 KM autoimmune pulmonary inflammation; cancer; tumour; gene therapy;
 KM SCR-1; supporting factor for the proliferation of stem cell.

XX Homo sapiens.

OS

PN WO200177169-A2.

XX

PD 18-OCT-2001.

XX

PF 05-APR-2001; 2001WO-US11208.

XX

PR 05-APR-2000; 2000US-0543774.

XX

PR 28-JUN-2000; 2000US-215733P.

PR 09-JAN-2001; 2001US-0757562.

XX

PR 05-FEB-2001; 2001US-266614P.

XX

PA (HYSE-) HYSEO INC.

PA (KIRI) KIRIN BEER KK.

XX

PI Tang TY, Labat I, Tillinghast JS, Sinku A, Liu C, Drmanac RT;

PI Stache-Crain B, Dickson M, Mize NK, Nishikawa M;

XX

DR MPI: 2001-657166/75.

DR N-PSDB: AAD21723.

XX

PT Novel stem cell growth factor like polypeptides and polynucleotides for
 PT identifying modulators useful for treating diseases such as Alzheimer's
 PT disease, cancer, rheumatoid arthritis, osteoporosis

PS Example 2; Page 166; 232pp; English.

XX

XX The patent discloses novel stem cell growth factor-like proteins and
 CC polynucleotides encoding them. Proteins of the invention are also known
 CC as supporting factor for the proliferation of stem cells (SCR-1). Stem
 CC cell growth factor-like proteins are useful for supporting proliferation
 CC or survival of a stem cell or germ cell which is preferably primordial
 CC cell, haematopoietic progenitor cell, pluripotent cell or totipotent
 CC cell. The haematopoietic progenitor cell cultured using stem cell
 CC growth factor-like proteins can replace as a graft for the bone marrow
 CC transplantation or cord blood transplantation for treating a variety
 CC of diseases such as immunodeficiency syndrome, agammaglobulinemia,
 CC Miskott-Aldrich syndrome, acquired immune deficiency syndrome (AIDS),
 CC thalassemia, haemolytic anaemia due to enzyme defect, congenital
 CC anaemia such as sickle cell anaemia, Gaucher's disease, lysosomal
 CC storage diseases such as mucopolysaccharidosis, adrenal white matter
 CC degeneration, a variety of cancer and tumours. Proteins of the
 CC invention are useful for treating diseases such as Parkinson's
 CC disease, Alzheimer's disease and other neurodegenerative diseases,
 CC thrombocytopaenia, immune deficiencies and disorders such as severe
 CC combined immunodeficiency (SCID) and autoimmune disorders such as
 CC multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis,
 CC and autoimmune pulmonary inflammation. Sequences of the invention are
 CC also useful in gene therapy. The present sequence is stem cell growth
 CC factor-like protein from human.

XX Sequence 160 AA;

S2

Query Match 59.6%; Score 903; DB 22; Length 160;
 Best Local Similarity 100.0%; Pred. No. 5,8e-64;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 CTKKADCDPCFKMNCCTCKSGFLHIGKLDNCPGGLFANNHTECVSYHCEVSENN 155
 DB 1 CTKKADCDPCFKMNCCTCKSGFLHIGKLDNCPGGLFANNHTECVSYHCEVSENN 60

QY 156 FMSPTCKKGTCTGFKGTETRVREIIHQPSAKGNLCPTNETKCTVQRKKCKGERGK 215
 DB 61 FMSPTCKKGTCTGFKGTETRVREIIHQPSAKGNLCPTNETKCTVQRKKCKGERGK 120

QY 216 GRERKKPKNGESKEAIPDSKLSSEKPEORENKKQ 255
 DB 121 GRERKKPKNGESKEAIPDSKLSSEKPEORENKKQ 160

RESULT 14

AB11374

XX ID AB11374 standard; peptide; 160 AA.

XX

AC AB11374;

XX

DT 11-JAN-2002 (first entry)

XX

DE Human secreted protein homologue, SEQ ID NO:1744.

XX

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KM haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KM inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KM proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KM myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KM chronic inflammatory condition; proliferative retinopathy;
 KM atherosclerosis; coronary heart disease; arterial ischaemia;
 KM bone disorder; osteoporosis; vascular growth disorder;
 KM tissue regeneration; wound healing; infection; immune disorder;
 KM cell culture; drug screening; gene therapy; antiinflammatory;
 KM antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KM cytosolic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KM antifungal; vulnery; antulcer.

XX

OS Homo sapiens.

XX

XX WO200157188-A2.

XX

PD 09-AUG-2001.

XX

PF 05-FEB-2001; 2001WO-US03800.

XX

PR 03-FEB-2000; 2000US-0496914.

XX

PR 27-APR-2000; 2000US-0560875.

XX

PA (HYSE-) HYSEO INC.

XX

PI Tang YT, Liu C, Drmanac RT;

PI MPI: 2001-457740/49.

XX

DR N-PSDB: ABA08618.

XX

PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer.

XX

XX Claim 20; Page 174-175; 1963pp; English.

XX

XX Sequences AB110981-AB112330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the

CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoietic regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness.
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.
 CC
 CC Sequence 160 AA;
 CC
 CC Query Match 59.6%; Score 903; DB 22; Length 160;
 CC Best Local Similarity 100.0%; Pred. No. 5.8e-64;
 CC Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 96 CTCKACADCDTCFNNKFTCKCKSGFYHLGKCLDNCPEGLFANNHTMECVSIHCEVSEMN 155
 CC DB 1 CTCKACADCDTCFNNKFTCKCKSGFYHLGKCLDNCPEGLFANNHTMECVSIHCEVSEMN 60
 CC QY 156 PWSPTCKKGTGCFKRGCTETRVREIIQHPSAKGNLCPTNETRCKCTVORRKCCKGGERGKX 215
 CC DB 61 PWSPTCKKGTGCFKRGCTETRVREIIQHPSAKGNLCPTNETRCKCTVORRKCCKGGERGKX 120
 CC QY 216 GREKRRKPPKNGESKEAIPDSKSLSSKEIPEORENKQOO 255
 CC DB 121 GREKRRKPPKNGESKEAIPDSKSLSSKEIPEORENKQOO 160
 CC
 CC RESULT 15
 CC AAM79312
 CC ID AAM79312 standard; Protein; 160 AA.
 CC XX
 CC AC AAM79312;
 CC XX
 CC DT 06-NOV-2001 (first entry)
 CC XX
 CC DE Human protein SEQ ID NO 2958.
 CC XX
 CC XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 CC KM vaccine; peptide therapy; stem cell growth factor; haematopoietic;
 CC KM tissue growth factor; immunomodulatory; cancer; leukaemia;
 CC KM nervous system disorder; arthritis; inflammation.
 CC XX
 CC OS Homo sapiens.
 CC XX
 CC XX WO200157190-A2.
 CC XX
 CC XX PD 09-AUG-2001.
 CC XX
 CC PF 05-FEB-2001; 2001MO-US04098.
 CC XX

PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-064936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 XX (HSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Dzmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 XX Xue AJ, Yang Y, Wejhtman T, Goodrich R;
 XX WPI: 2001-476283/51.
 XX N-PSDB; AAK52445.
 XX
 XX Nucleic acids encoding polypeptides with cytokine-like activities,
 XX useful in diagnosis and gene therapy -
 XX
 XX Claim 20; Page 214; 6221pp; English.
 XX
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 XX encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 XX cytokine, cell proliferation or cell differentiation or which may induce
 XX production of other cytokines in other cell populations. The
 XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
 XX peptide therapy. The polypeptides have various cytokine-like activities,
 XX e.g. stem cell growth factor activity, haematopoietic regulatory
 XX activity, tissue growth factor activity, immunomodulatory activity and
 XX activin/inhibin activity and may be useful in the diagnosis and/or
 XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
 XX inflammation.
 XX Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 XX (AAM80020) are omitted as the relevant pages from the sequence listing
 XX were missing at the time of publication.
 XX
 XX Sequence 160 AA;
 XX
 XX Query Match 59.6%; Score 903; DB 22; Length 160;
 XX Best Local Similarity 100.0%; Pred. No. 5.8e-64;
 XX Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 96 CTCKACADCDTCFNNKFTCKCKSGFYHLGKCLDNCPEGLFANNHTMECVSIHCEVSEMN 155
 XX DB 1 CTCKACADCDTCFNNKFTCKCKSGFYHLGKCLDNCPEGLFANNHTMECVSIHCEVSEMN 60
 XX QY 156 PWSPTCKKGTGCFKRGCTETRVREIIQHPSAKGNLCPTNETRCKCTVORRKCCKGGERGKX 215
 XX DB 61 PWSPTCKKGTGCFKRGCTETRVREIIQHPSAKGNLCPTNETRCKCTVORRKCCKGGERGKX 120
 XX QY 216 GREKRRKPPKNGESKEAIPDSKSLSSKEIPEORENKQOO 255
 XX DB 121 GREKRRKPPKNGESKEAIPDSKSLSSKEIPEORENKQOO 160
 XX

Search completed: May 6, 2003, 14:51:39
 Job time : 38 secs

RT morphogenetic proteins at distinct sites during embryogenesis.";
RL J. Cell Biol. 134:181-191 (1996).
RN (6)
RP DEVELOPMENTAL EXPRESSION
RX MEDLINE=97436919; PubMed=9291583;
RA Rancourt S.L., Rancourt D.E.;
RT "Murine subtilisin-like proteinase SPC6 is expressed during embryonic
implantation, somitogenesis, and skeletal formation.";
RL Dev. Genet. 21:75-81(1997).
CC -1- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPROTEASE ACTIVITY
WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE
OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF. MAY BE INVOLVED
FOR THE MATURATION OF GASTROINTESTINAL PEPTIDES. MAY BE INVOLVED
IN THE CELLULAR PROLIFERATION OF ADRENAL CORTEX VIA THE ACTIVATION
OF GROWTH FACTORS.
CC -1- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
PROTEIN BY CLEAVAGE OF ARG-XAA-YAA-ARG-|-ZAA BONDS, WHERE XAA
CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
CC SUBCELLULAR LOCATION: PCSA IS SECRETED THROUGH THE REGULATED
SECRETORY PATHWAY. PCSB IS A TYPE I MEMBRANE PROTEIN LOCALIZED TO
A PARANUCLEAR POST-GOLGI NETWORK COMPARTMENT IN COMMUNICATION WITH
EARLY ENDOPLASM.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS: PCSB/LONG (SHOWN HERE)
AND PCSA/SHORT: ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: PCSA IS EXPRESSED IN MOST TISSUES, BUT IS MOST
ABUNDANT IN THE INTESTINE AND ADRENALS. PCSB IS EXPRESSED IN THE
INTESTINE, ADRENALS AND LUNG BUT NOT IN THE BRAIN.
CC -1- DEVELOPMENTAL STAGE: WEAKLY EXPRESSED THROUGHOUT THE EMBRYO,
EXCEPT IN THE DEVELOPING NERVOUS SYSTEM, THE RIBS AND THE LIVER,
BUT MARKEDLY UPREGULATED AT DISCRETE SITES DURING DEVELOPMENT. AT
E6.5, PROMINENT EXPRESSION OBSERVED IN DIFFERENTIATED DECIDUA. AT
E7.5, INTENSE EXPRESSION IN EXTRAMERIONIC ENDODERM, AMNION AND
NASCENT MESODERM. AT 8.5, ABUNDANT EXPRESSION IN SOMITES AND YOLK
SAC FOLLOWED BY A CONFINATION TO DERMATOTOME COMPARTMENT. BETWEEN
E9.5 AND E11.5, ABUNDANT EXPRESSION IN AER (THICKENED ECTODERMAL
CELLS OF LIMB BUDS). AT E12.5, EXPRESSION IN THE LIMBS IS CONFINED
TO THE CONDENSING MESENCHYM SURROUNDING THE CARILAGE. AT THIS
STAGE, STRONG EXPRESSION ALSO DETECTED IN VERTEBRAL AND FACIAL
CARILAGE PRIMORDIA AND IN THE MUSCLE OF THE TONGUE. AT E16.5,
ABUNDANT EXPRESSION IN EPITHELIAL CELLS OF THE INTESTINAL VILLI.
CC ISOFORM A IS MOST ABUNDANT AT ALL STAGES BUT SIGNIFICANT LEVELS OF
ISOFORM B OCCUR AT E12.5.
CC -1- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE
ASSISTING THE FOLDING OF THE ZMOGEN WITHIN THE ENDOPLASMIC
RETICULUM.
CC -1- DOMAIN: AC 1 AND AC 2 (CLUSTERS OF ACIDIC AMINO ACIDS) CONTAIN
SORTING INFORMATION. AC 1 DIRECTS TGN LOCALIZATION AND INTERACTS
WITH THE TGN SORTING PROTEIN PACS-1.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -1- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entitled requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D17583; BAA04507.1; -;
DR EMBL; D12619; BAA02143.1; -;
DR EMBL; I14932; AAT74636.1; -;
DR PIR; JX0248; JX0248.
DR PIR; A48225; A48225.
DR HSSP; Q99405; IMPT.
DR MEROPS; S08.076; -;
DR MGD; MG1:97515; Pcsk5.
DR InterPro; IPR000563; EGF-like.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR002884; P_domain.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF01483; P; PARTIAL.

DR PRINTS; PR00723; SUBTILISIN.
DR ProDom; PD000717; P domain; 1.
DR SMART; SM00181; EGF_3.
DR SMART; SM00001; EGF-like; 2.
DR SMART; SM00261; FU_-22.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
KM Cleavage on pair of basic residues; Repeat; Alternative splicing;
KN Transmembrane.
FT SIGNAL 1 34
FT PROPEP 35 116
FT CHAIN 117 1877
FT
FT DOMAIN 117 1768
FT TRANSMEM 1769 1789
FT DOMAIN 1790 1877
FT DOMAIN 117 452
FT DOMAIN 464 602
FT DOMAIN 638 1753
FT DOMAIN 1825 1844
FT DOMAIN 1856 1877
FT SITE 116 117
FT SITE 521 523
FT ACT_SITE 173 173
FT ACT_SITE 214 214
FT ACT_SITE 388 388
FT CARBOHYD 227 227
FT CARBOHYD 383 383
FT CARBOHYD 667 667
FT CARBOHYD 754 754
FT CARBOHYD 804 804
FT CARBOHYD 854 854
FT CARBOHYD 951 951
FT CARBOHYD 1016 1016
FT CARBOHYD 1220 1220
FT CARBOHYD 1317 1317
FT CARBOHYD 1523 1523
FT CARBOHYD 1711 1711
FT CARBOHYD 1733 1733
FT VARSPPLIC 878 915
FT
FT VARSPPLIC 916 1877
FT VARSPPLIC 209287 MW; EC850E2DF20EALC3 CRC64;
SQ SEQUENCE 1877 AA; 209287 MW; EC850E2DF20EALC3 CRC64;
Query Match 12.3%; Score 186.5; DB 1; Length 1877;
Best Local Similarity 25.0%; Pred. No. 7.3e-06;
Matches 66; Conservative 31; Mismatches 90; Indels 77; Gaps 15;
QY 39 SGGCGGCGATCSGYNQ--CLGCKPLFALERIGMOIGVGLSSCPSTGYGRYPD--N 94
DB 1209 NQPHSSCKTC--NSGLCASCPGTOMYLMD-----ACVSPCGG---TWSVTSG 1253
QY 95 KCTCKKADCDCTCFNKFCCTCKS-----GFYLHLGCKLIDNCPGLLEANNHTME-CVSIYHC 149
DB 1254 SCNCSSDVCVSGGADLCQCGCLSPDNTLLHBCRCVHSPBGFYAKDGVCHCSS--PC 1311
QY 150 EYSEMNPMWSPCTKKG-----KTC-----GFKRGTRRRELIQHPSAKGML- 190
DB 1312 KTCGNATSCNCGSDGFVDHGVCKKTCPEKGVAVBVCVCHGCEPRQDDILHKTCKEMP 1371
QY 191 -----CPPT--NETRCKTVORRKKCKGKRGKRGKRRKKPKGSKKALPDS 236
DB 1372 DFLYNDMCHRSCPKSFYDMDQVCPCHKN-----LECNQPFEDDKVCACADTS 1420
QY 237 KSLSS---KEIPE-----ORENKO 253
DB 1421 KALNGLCLDECEBGTGKEEENDE 1444
RESULT 2

PKCS_BRACL
ID PKCS_BRACL STANDARD: PRT: 1696 AA.
AC 09N15; 09N16; 09N14;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE (Protein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-))
GN PC6.
OS Branchiostoma californiensis (California lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.
NCBI_TaxID=7738;
RX SEQUENCE FROM N.A. (ISOFORMS A, B AND C).
RA MEDLINE=20175281; PubMed=10708688;
OLIVA A.A. Jr., Chan S.J., Steiner D.F.;
RT "Evolution of the prohormone convertases: identification of a
RT homologue of PC6 in the protochordate amphioxus";
RL Biochim. Biophys. Acta 1477:338-348(2000).
CC -I- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPEPTIDASE ACTIVITY
CC WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE
CC OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF (BY SIMILARITY).
CC -I- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
CC PROTEIN CLEAVAGE OF ARG-XAA-YAA-ARG-|-ZAA BONDS, WHERE XAA
CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
CC -I- SUBCELLULAR LOCATION: ISOFORM A AND ISOFORM C ARE SECRETED.
CC ISOFORM B IS A TYPE I MEMBRANE PROTEIN.
CC -I- ALTERNATIVE PRODUCTS: 3 ISOFORMS, A, B (SHOWN HERE) AND C; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -I- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE
CC ASSISTING THE FOLDING OF THE ZMOGEN WITHIN THE ENDOPLASMIC
CC RETICULUM.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY 58.
CC -I- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF184615; AAF26300.1; -
DR EMBL; AF184616; AAF26301.1; -
DR EMBL; AF184617; AAF26302.1; -
DR HSP; O99405; 1MPT.
DR MEROPS; S08.028; -
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR002884; P_domain.
DR InterPro; IPR002029; Peptidase_S8.
DR Pfam; PF01483; P; 1.
DR Pfam; PF01483; P; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PRODOM; PD000717; P_domain; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00261; FU; 17.
DR PROSITE; PS00136; SUBTILASIN ASP; FALSE_NEG.
DR PROSITE; PS00137; SUBTILASIN HIS; 1.
DR PROSITE; PS00138; SUBTILASIN SER; 1.
KW Hydrolyase: Serine protease; Glycoprotein; Zymogen; Signal;
KW Cleavage on pair of basic residues; Repeat; Alternative splicing;
KW Transmembrane.
FT SIGNAL 1 25 POTENTIAL.
FT PROPEP 26 110 POTENTIAL.
FT CHAIN 111 1696 PROPEPTIDE CONVERTASE SUBTILISIN/KEXIN
FT TYPE 5.
FT DOMAIN 111 1618 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1619 1639 POTENTIAL.
FT DOMAIN 1640 1696 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 111 488 CATALYTIC.

FT DOMAIN 496 637 HOMO B.
FT SITE 664 1649 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT ACT_SITE 110 111 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 192 192 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 233 233 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 407 407 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 246 246 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 529 529 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 885 885 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARSPLIC 1259 1323 DTTIDRGECITSCGPGKMGREKKAACHPTCKECSDEY
FT DTTIDRGECITSCGPGKMGREKKAACHPTCKECSDEY
FT PREVSILAEIAGHRLRSLTPPOSIPDVTGLGDRRL
FT TTTSAGRCA (IN ISOFORM C).
FT MISSING (IN ISOFORM C).
FT CHPTCKECSDEYDDTCTACNDGFLTLTDASSCEAGCPGQFL
FT HHGDCDSCHRECKTC -> IARCVDRDRSRMCDLVRFNFC
FT VRRYFVRCCTGCKLYMEDRRMRGSSOPTOGRN (IN
FT ISOFORM A).
FT VARSPLIC 1344 1696 MISSING (IN ISOFORM A).
FT SEQUENCE 1696 AA; 188410 MM; 281CBE1784257C8D CRC64;
SO QUERY MATCH 12.24; Score 184.5; DB 1; Length 1696;
Best Local Similarity 22.74; Pred. No. 9.2e-06;
Matches 64; Conservative 25; Mismatches 74; Indels 119; Gaps 13;

OY 34 MHPNVSQGGGCGCATCS--DYNGCLSCRPFLF----- 64
DB 1328 LHHGDDSDCHRECKTCDDPHNDLSCQPSGYINDQCSHCPSTFEYEDDSGETVL 1387
OY 65 -----ALERTGMKQIGVCLSSCPSPGYTRYPPDI-NKCTKCYA 101
DB 1388 QCLCHVNCCTCHGEGEDMECANDIKYKQDRCVTEGCEGH--YPLDTMCCOCWS 1443
OY 102 DCDTCF--NKNFCTCKSGFYHLGKLDNCEGL-----EANN 138
DB 1444 DCTCGPRNDQCVTCPPNYLVLGKLEDCPEGYDTPMROKEGEGCHSCATCEGNN 1503
OY 139 HTMECHSYH-----CEVSEW-----NMSPTCKKATCGFKGTERREI 181
DB 1504 Y--NCLSCPSYSGKLAGCVPCMBEHYVETKQICEBDNSCTC--KSTADHCLSC 1558
OY 182 QHP--SAKGNLC-----PTNE-----TRKCTVYORK 205
DB 1559 EARYGYHMKHLCTACCEGSPENEYCCICHESTRLCTIDRE 1600

RESULT 3
FUR2_DROME STANDARD: PRT: 1680 AA.
ID FUR2_DROME
AC P30432;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Furin-like protease 2 precursor (EC 3.4.21.75) (Furin 2).
GN FUR2.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=923181036; PubMed=1512259;
RA Roebroek A.J.M., Greeners J.W.M., Pauli I.G.L., Kurzik-Dunke U.,
RA Roebroek A.J.M., Galleff E.A.F., Leunissen J.A.M., van de Ven W.J.,
RT "Cloning and functional expression of Drosophila melanogaster with multiple
RT propeptide processing enzyme of Drosophila melanogaster with multiple
RT repeats of a cysteine motif.";
RL J. Biol. Chem. 267:17208-17215(1992).
CC -I- FUNCTION: FURIN IS LIKELY TO REPRESENT THE UBIQUITOUS ENDOPEPTIDASE
CC ACTIVITY WITHIN CONSTITUTIVE SECRETORY PATHWAYS AND CAPABLE OF
CC CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF (BY SIMILARITY).
CC -I- CATALYTIC ACTIVITY: Release of mature proteins from their

CC proproteins by cleavage of Arg-Xaa-Yaa-Arg|-Zaa bonds, where Xaa
 CC can be any amino acid and Yaa is Arg or Lys. Releases albumin,
 CC complement component C3 and von Willebrand factor from their
 CC respective precursors.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8. FURIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M94375.1, AAA28551.1, -
 DR PIR: A43434, A43434.
 DR HSSP: Q99405, IMP1.
 DR MEROPS: S08.049, -
 DR FLYBASE: FBgn0004598, Fur2.
 DR InterPro: IPR000561, EGF-like.
 DR InterPro: IPR002174, Furin-like.
 DR InterPro: IPR002884, P domain.
 DR InterPro: IPR000209, Peptidase_S8.
 DR Pfam: PF00082, Peptidase_S8_1.
 DR Pfam: PF01483, P, PARTIAL.
 DR PRINTS: PR00723, SUBTILISIN.
 DR ProDom: PD000717, P_domain, 1.
 DR SMART: SM00181, EGF, 1.
 DR SMART: SM00261, FU, 10.
 DR PROSITE: PS00136, SUBTILISIN ASP, 1.
 DR PROSITE: PS00137, SUBTILISIN HIS, 1.
 DR PROSITE: PS00138, SUBTILISIN SER, 1.
 DR KX Hydroxylase; Serine protease; Glycoprotein; Signal; Transmembrane;
 KW Multigene family; Zymogen; Repeat
 FT SIGNAL 1 ?
 FT PROPEP 1 319
 FT CHAIN 320 1680
 FT ACT_SITE 418 418
 FT ACT_SITE 457 457
 FT ACT_SITE 638 638
 FT DOMAIN 962 1444
 FT REPEAT 962 1007
 FT REPEAT 1008 1057
 FT REPEAT 1058 1104
 FT REPEAT 1105 1153
 FT REPEAT 1154 1205
 FT REPEAT 1206 1254
 FT REPEAT 1255 1299
 FT REPEAT 1300 1346
 FT REPEAT 1347 1393
 FT REPEAT 1394 1444
 FT TRANSMEM 1508 1532
 FT DOMAIN 1533 1680
 FT CARBOHD 3
 FT CARBOHD 109 3
 FT CARBOHD 130 130
 FT CARBOHD 203 203
 FT CARBOHD 443 443
 FT CARBOHD 481 481
 FT CARBOHD 928 928
 FT CARBOHD 1061 1061
 FT CARBOHD 1182 1182
 FT CARBOHD 1273 1273
 FT CARBOHD 1278 1278
 FT CARBOHD 1440 1440
 SQ SEQUENCE 1680 AA; 183599 MW; 0A99CEB770A8E293 CRC64;
 Query Match 11.6%; Score 176.5; DB 1; Length 1680;
 Best Local Similarity 28.0%; Pred. No. 3,3e-05;
 Matches 60; Conservative 24; Mismatches 77; Indels 53; Gaps 12;
 3 HURLSWIFILNMEYISGNASRRGRMRHNPVSGCGGCGATCSGY-NGLCSCKPR 61

Db 1038 HLHVID-LAVCLQCPDGYFENS---RNRTCP-----CEPNACSCDHPDYCTSCDH 1087
 Qy 62 LFFALRIQMKIGVCLSSCPSSGYGTRYPDINKCTKCADDTCF--NKNFCTCKSGF 119
 Db 1088 LVMHEK-----CYSACPLDYET---EDNKAFCFHCSTCATCNGPTDDCITCSSR 1136
 Qy 120 YLHGLCLNCGEGELANNHMECVSTVHCSESENNPMSPTCKKCTCGFGRGTEPV-R 178
 Db 1137 YAMONCLISCDGEGYAKKRLBGM-----PCDEGCKTC-----TSGVCS 1177
 Qy 179 EIIQHSAGNLCPPNETRKYCTQVRKK-CQKGE 211
 Db 1178 ECLQNT-----LNKRDKCIYSGSGCSESE 1203
 RESULT 4
 PACE4_HUMAN STANDARD; PRT; 969 AA.
 AC P9122; Q15099; Q15100; Q9UEU1; Q9UEU2; Q9UEU7; Q9UEU8; Q9UEU9;
 AC Q9UEG7; Q9Y4G9; Q9Y4H0; Q9Y4H1;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Paired basic amino acid cleaving enzyme 4 precursor (EC 3.4.21.-)
 DE (subtilisin/kexin-like protease PACE4) (subtilisin-like propeptin
 DE convertase 4) (SPC4).
 GN PACE4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS PACE4A-I AND PACE4B).
 RC TISSUE=Hepatosoma, and kidney;
 RX MEDLINE=92075167; PubMed=1741956;
 RA Kiefer M.C., Tucker J.E., Jon R., Landsberg K.E., Saltman D.,
 RA Barr P.J.;
 RT "Identification of a second human subtilisin-like protease gene in
 RT the tes/fps region of chromosome 15.";
 RL DNA Cell Biol. 10:757-769(1991).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS PACE4C AND PACE4D).
 RC TISSUE=Placenta;
 RX MEDLINE=94235049; PubMed=8179631;
 RA Tsuji A., Higashine K., Hine C., Mori K., Tamai Y., Nagamune H.,
 RA Matsuda Y.;
 RT "Identification of novel cDNAs encoding human kexin-like protease,
 RT PACE4 isoforms.";
 RL Biochem. Biophys. Res. Commun. 200:943-950(1994).
 RN [3]
 RP ERRATUM.
 RX MEDLINE=95071480; PubMed=7980617;
 RA Tsuji A., Higashine K., Hine C., Mori K., Tamai Y., Nagamune H.,
 RA Matsuda Y.;
 RT "Identification of novel cDNAs encoding human kexin-like protease,
 RT PACE4 isoforms.";
 RL Biochem. Biophys. Res. Commun. 204:1381-1382(1994).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM PACE4A-II).
 RC TISSUE=Placenta;
 RX Mori K., Imamaki A., Kii S., Nagamune H., Nagahama M., Tsuji A.,
 RA Matsuda Y.;
 RT "Identification of a novel PACE4 isoform, PACE4E.";
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORMS PACE4E-I AND PACE4E-II).
 RC TISSUE=Cerebellum;
 RX MEDLINE=97135942; PubMed=9192737;
 RA Mori K., Kii S., Tsuji A., Nagahama M., Imamaki A., Hayashi K.,
 RA Akamatsu T., Nagamune H., Matsuda Y.;
 RT "A novel human PACE4 isoform, PACE4E is an active processing protease
 RT containing a hydrophobic cluster at the carboxy terminus.";
 RL J. Biochem. 121:941-948(1997).

RN [6]
RP SEQUENCE FROM N.A. (ISOFORMS PACE4A-I; A-II; CS; D; E-I; E-II).
RX MEDLINE=99021085; PubMed=9378725;
RA Tsuji A., Hine C., Tamai Y., Tomemoto K., Mori K., Yoshida S.,
Bando M., Sakai E., Mori K., Akamatsu T., Matsuda Y.;
RT "Genomic organization and alternative splicing of human PACE4 (SPC4),
RT kexin-like processing endoprotease.";
RL J. Biochem. 122:438-452(1997).
RN [7]
RP ALTERNATIVE SPLICING (ISOFORM PACE4CS).
RX MEDLINE=97064242; PubMed=8906861;
RA Zhong M., Benjamet S., Lazure C., Munzer S., Seidah N.G.;
RT "Functional analysis of human PACE4-A and PACE4-C isoforms:
RT identification of a new PACE4-CS isoform.";
RL FEBS Lett. 396:31-36(1996).
RN [8]
RP CHARACTERIZATION.
RX MEDLINE=99233559; PubMed=10215603;
RA Susic J.F., Moehring J.M., Innocencio N.M., Luchini J.W.,
RT "Endoprotease PACE4 is Ca²⁺-dependent and temperature-sensitive and
RT can partly rescue the phenotype of a furin-deficient cell strain.";
RL Biochem. J. 339:639-647(1999).
RN [9]
RP PROCESSING.
RX MEDLINE=98408849; PubMed=9738469;
RA Nagahama M., Taniguchi T., Hashimoto E., Imamaki A., Mori K.,
RA Tsuji A., Matsuda Y.;
RT "Biosynthetic processing and quaternary interactions of proprotein
RT convertase SPC4 (PACE4).";
RL FEBS Lett. 444:155-159(1998).
CC -I- FUNCTION: LIKELY TO REPRESENT AN ENDOPEPTIDASE ACTIVITY WITHIN THE
CC CONSTITUTIVE SECRETORY PATHWAY, WITH UNIQUE RESTRICTED
CC DISTRIBUTION IN BOTH NEUROENDOCRINE AND NON-NEUROENDOCRINE TISSUES
CC AND CAPABLE OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF.
CC -I- CATALYTIC ACTIVITY: RELEASE OF NATURE PROTEINS FROM THEIR
CC PROTEINETS BY CLEAVAGE OF ARG-XAA-YAA-ARG-|-ZAA BONDS,
CC WHERE XAA CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
CC -I- COFACTOR: PACE4A IS PROBABLY CALCIUM-DEPENDENT.
CC -I- SUBUNIT: THE PACE4A-I PRECURSOR PROTEIN SEEMS TO EXIST IN THE
CC RETICULUM ENDOPLASMIC AS BOTH A MONOMER AND A DIMER-SIZED COMPLEX
CC WHEREAS MATURE PACE4A-I EXISTS ONLY AS A MONOMER, SUGGESTING THAT
CC PROPEPTIDE CLEAVAGE AFFECTS ITS TERTIARY OR QUATERNARY STRUCTURE.
CC -I- SUBCELLULAR LOCATION: PACE4A-I AND PACE4A-II ARE SECRETED. PACE4C
CC AND PACE4CS ARE NOT SECRETED AND REMAIN PROBABLY IN ZYMOGEN FORM
CC IN ENDOPLASMIC RETICULUM. PACE4E-I AND PACE4E-II ARE RETAINED
CC INTRACELLULARLY PROBABLY THROUGH A HYDROPHOBIC CLUSTER IN THEIR C-
CC TERMINUS. PACE4B MIGHT BE SECRETED.
CC -I- ALTERNATIVE PRODUCTS: 8 ISOFORMS: PACE4A-I/PACE4 (SHOWN HERE),
CC PACE4A-II, PACE4B/PACE4.1, PACE4C, PACE4CS, PACE4D, PACE4E-I AND
CC PACE4E-II; ARE PRODUCED BY ALTERNATIVE SPLICING. ISOFORMS PACE4B,
CC C, CS AND D MIGHT BE ENZYMATICALLY INACTIVE.
CC -I- TISSUE SPECIFICITY: EACH PACE4 ISOFORM EXHIBITS A UNIQUE
CC RESTRICTED DISTRIBUTION. PACE4A-I IS EXPRESSED IN HEART, BRAIN,
CC PLACENTA, LUNG, SKELETAL MUSCLE, KIDNEY, PANCREAS, BUT AT
CC COMPARTMENTALLY HIGHER LEVELS IN THE LIVER. PACE4A-II IS AT LEAST
CC EXPRESSED IN PLACENTA. PACE4B WAS ONLY FOUND IN THE EMBRYONIC
CC KIDNEY CELL LINE FROM WHICH IT WAS ISOLATED. PACE4C AND PACE4D ARE
CC EXPRESSED IN PLACENTA. PACE4E-I IS EXPRESSED IN CEREBELLUM,
CC PLACENTA AND PITUITARY. PACE4E-II IS AT LEAST PRESENT IN
CC CEREBELLUM.
CC -I- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE
CC ASSISTING THE FOLDING OF THE ZYMOGEN WITHIN THE ENDOPLASMIC
CC RETICULUM. ISOFORM PACE4D LACKS THE PROPEPTIDE DOMAIN.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -I- SIMILARITY: CONTAINS 1 HOWO B/P DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; M80482; AAA5998.1; JOINED
DR EMBL; AB001914; BAA21620.1; JOINED
DR EMBL; AB001988; BAA21620.1; JOINED
DR EMBL; AB001900; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB001909; BAA21620.1; JOINED
DR EMBL; AB001910; BAA21620.1; JOINED
DR EMBL; AB001901; BAA21620.1; JOINED
DR EMBL; AB001902; BAA21620.1; JOINED
DR EMBL; AB001903; BAA21620.1; JOINED
DR EMBL; AB001904; BAA21620.1; JOINED
DR EMBL; AB001905; BAA21620.1; JOINED
DR EMBL; AB001906; BAA21620.1; JOINED
DR EMBL; AB001907; BAA21620.1; JOINED
DR EMBL; AB001908; BAA21620.1; JOINED
DR EMBL; AB

[illegible]

```

FT CARBOHYD 804 804 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 854 854 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1710 1710 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1732 1732 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 878 915 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX ->
FT VARSPLIC 878 915 AAESEMAEGFCMVLVKKNNLCCORRVLAQOLCCCTCFPOG
FT VARSPLIC 878 915 (IN ISOFORM PCSA).
FT VARSPLIC 878 915 MISSING (IN ISOFORM PCSA).
SQ SEQUENCE 1877 AA; 20788 MW; 890955DC0534444 CRC64;

Query Match 11.1%; Score 168; DB 1; Length 1877;
Best Local Similarity 26.3%; Pred. No. 0.0001;
Matches 49; Conservative 25; Mismatches 54; Indels 58; Gaps 11;

OY 36 PNVSG-GCGG-----GCATCSYNGCLSCPKRLFPALERIGKMGVGLSCSPSYGYTRY 90
DB 640 PECSEVCGDGPDPHCTDLHYHKLKNNTR-----ICVSSCPGPHF---H 682
OY 91 PINKCTCKKADCDPTCNKXF--CTCKSGPYLH--LGKCLDNPBGLANHTMECVSI 146
DB 683 ADKRCRCACAPNCSCGSHADCLCKYGLNETSSCVAOCPESSYODIKNIC--- 739
OY 147 VHCSEVSEWNPSPCTKKKTC-GFKRGTRVRELIHPSAKGNLCPTNETRKTCTVOR 205
DB 740 -----GKCSNCKTCTGFHNCTE-----CKGGL---SLOGSRCSV--- 771
OY 206 KCOGE 211
DB 772 TCEDCG 777

RESULT 6
PKCS_HUMAN
ID PKCS_HUMAN STANDARD; PRT; 913 AA.
AC 092824; Q13527;
DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE Proprotein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)
DE (Protein convertase PCS) (Subtilisin/kexin-like protease PCS)
DE (Convertase PCS) (PC6) (hPC6).
GN PCSK5 OR PCS OR PC6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell;
RA MEDLINE=96353880; Pubmed=8755538;
RA Miranda L., Wolf J., Pichuanes S., Duke R., Franzusoff A.;
RT "Isolation of the human PC6 gene encoding the putative host protease
RT for HIV-1 gp160 processing in CD4+ T lymphocytes."
RL Proc. Natl. Acad. Sci. U.S.A. 93:7695-7700(1996).
RN [2]
RP REVISIONS.
RA Franzusoff A., Miranda L., Wolf J., Pichuanes S., Lu Y., Duke R.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 15-913 FROM N.A.
RA Ruedelhuber T.L.;
CC -1- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPEPTIDASE ACTIVITY
CC WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE
CC OF CLEAVAGE AT THE RX(R/R)R CONSENSUS MOTIF.
CC -1- CATALYTIC ACTIVITY: RELEASE OF MAJORE PROTEINS FROM THEIR
CC PROPEPTIDES BY CLEAVAGE OF ARG-XAA-YAA-ARG-1-ZAA BONDS, WHERE XAA
CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
CC -1- SUBCELLULAR LOCATION: Secreted (by similarity).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS SEEMS TO BE PRODUCED BY
CC ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN T-LYMPHOCYTES.
CC -1- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE

```

```

CC ASSISTING THE FOLDING OF THE ZMOGEN WITHIN THE ENDOPLASMIC
CC RETICULUM.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -1- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sid.ch).
CC -----
OY 21 GSNASRGGRORRMHPNVSGCGCATCSYNG--CLSCPKRLFPALERIGKMGVGL 78
DB 677 GHYHDK-KRCCKAPN-----CESFGSHGQCMSCYGFLL-----NEINSCV 721
OY 79 SSCPSG-YGGRYPDINKCTKACDDTFCNPKCKSGFYHLKCLDNPBGLAN 137
DB 722 THCPDSYODTKK--NLCKRCSENCCTCFHNCTEGRDLSSLOGSRCSVSEDRYEN 778
OY 138 NHTME-----CVS-----YHCVSEW-----NPSWSPCTCK 163
DB 779 GQDCPCRFATCAGAGAGCINCTEGRFMEGRCVQSGSISYFPHSENGYKCKKC 838
OY 164 GKTG-----GFKRGTE-----TRVRELIHPSAKGNLCPTNETRKTCTV 202

```

```

Db      839 DISCTGNGGFGFKCTSCPSGYLLDLGCMQGAICKQATESSNMEGGFCMLVKXNNLC- 896
Oy      203 ORKCKQK 209
Db      897 ORKVLQK 903

RESULT 7
WIFI_HUMAN
ID      16-OCT-2001 (Rel. 40, Created)
AC      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Wnt inhibitory factor 1 precursor (WIF-1).
GN      WIF1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;

(1)
RP      SEQUENCE FROM N.A.
RX      MEDLINE=99215557; PubMed=10201374;
RA      Hsieh J.-C., Kodjabachian L., Rebber M.L., Rattner A.,
RA      Smallwood P.M., Samos C.H., Nusse R., Dawid I.B., Nathans J.;
RT      "A new secreted protein that binds to Wnt proteins and inhibits their
RT      activities."
RL      Nature 398:431-436(1999).
CC      - FUNCTION: BINDS TO WNT PROTEINS AND INHIBITS THEIR ACTIVITIES. MAY
CC      BE INVOLVED IN MESODERM SEGMENTATION.
CC      - SUBCELLULAR LOCATION: Secreted.
CC      - SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; AF129222; AAD25402.1; -
DR      HSSP; P00743; ICDF.
DR      Genew; HGNC:18081; WIF1.
DR      MIM; 605186; -
DR      InterPro; IPR000561; EGF-like.
DR      InterPro; IPR002049; Laminin_EGF.
DR      InterPro; IPR003306; WIF.
DR      Pfam; PF00008; EGF; 5.
DR      Pfam; PF02019; WIF; 1.
DR      PRINTS; PR00011; EGF_LAMININ.
DR      SMART; SM00181; EGF; 5.
DR      SMART; SM00469; WIF; 1.
DR      PROSITE; PS00022; EGF_1; 5.
DR      PROSITE; PS01186; EGF_2; 4.
KM      Repeat: EGF-like domain; Signal: Developmental protein.
FT      SIGNAL 1 28
FT      CHAIN 29 379
FT      DOMAIN 177 208 EGF-LIKE 1.
FT      DOMAIN 209 240 EGF-LIKE 2.
FT      DOMAIN 241 272 EGF-LIKE 3.
FT      DOMAIN 273 304 EGF-LIKE 4.
FT      DOMAIN 305 336 EGF-LIKE 5.
FT      DISULFID 177 186 POTENTIAL.
FT      DISULFID 182 192 POTENTIAL.
FT      DISULFID 198 200 POTENTIAL.
FT      DISULFID 209 218 POTENTIAL.
FT      DISULFID 214 224 POTENTIAL.
FT      DISULFID 230 232 POTENTIAL.
FT      DISULFID 241 250 POTENTIAL.
FT      DISULFID 246 256 POTENTIAL.
FT      DISULFID 262 264 POTENTIAL.

```

```

FT      DISULFID 273 282 POTENTIAL.
FT      DISULFID 278 288 POTENTIAL.
FT      DISULFID 294 296 POTENTIAL.
FT      DISULFID 305 314 POTENTIAL.
FT      DISULFID 310 320 POTENTIAL.
FT      DISULFID 326 328 POTENTIAL.
FT      CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ      SEQUENCE 379 AA; 41512 MW; 27782370A266E784 CRC64;

Query Match 10.5%; Score 158.5; DB 1; Length 379;
Best Local Similarity 24.9%; Pred. No. 0.00017;
Matches 60; Conservative 26; Mismatches 72; Indels 83; Gaps 16;

Oy      42 CQGGCT--CSDVNGC-----LSKRLPFLERIGMKQIGVCS-----C 81
Db      182 CPGGNGGFCNERRICCPDPFGPHCEKALCTPCNNG-----GLCVPGFCIC 232
Oy      82 PSQVYTRPDINKTKC-KADCD-TCFNKNFCTKCKSGFYLLHKLCLDNCEGLNANN 139
Db      233 PPGFYG-----VNCCKXKNCSTTCFNGTC-----FY-PKCI--CPGLEGE-- 271
Oy      140 TMECVSVHCEVSEWNPSPCTKKGTCGFGKGTETRYRELIQHSAKGNLCRP----- 193
Db      272 -----QCEISKCP--QPCRNKGKIG--KSKCKSGKYQGDLSKPYCEPGCGANG 318
Oy      194 -TWETKCTVQRKCKCKGCKGKGRERK-----KPNKESKKAIPDSKSLSSKEIP 246
Db      319 TCEPRKCK-----QCEGSMHGRHNKRYEALIHLPAGQLRQHTPSLKAKKEERRDP 373
Oy      247 E 247
Db      374 E 374

RESULT 8
PAC4_RAT
ID      063415;
AC      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Paired basic amino acid cleaving enzyme 4 precursor (EC 3.4.21.-)
DE      (Subtilisin/kexin-like protease PACE4) (Subtilisin-like proprotein
DE      convertase 4) (SPC4).
GN      PACE4.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;

(1)
RP      SEQUENCE FROM N.A.
RX      STRAIN=Sprague-Dawley; TISSUE=Hypothalamus, and Pituitary;
RX      MEDLINE=94349873; PubMed=8070361;
RA      Johnson R.C., Darlington D.N., Hand T.A., Bloomquist B.T., Mains R.E.;
RT      "PACE4, a subtilisin-like endoprotease prevalent in the anterior
RT      pituitary and regulated by thyroid status."
RL      Endocrinology 135:1178-1185(1994).
CC      - FUNCTION: LIKELY TO REPRESENT AN ENDOPEPTASE ACTIVITY WITHIN THE
CC      CONSTITUTIVE SECRETORY PATHWAY. WITH UNIQUE RESTRICTED
CC      DISTRIBUTION IN BOTH NEUROENDOCRINE AND NON-NEUROENDOCRINE TISSUES
CC      AND CAPABLE OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF.
CC      - CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
CC      PROPEPTIDES BY CLEAVAGE OF ARG-XAA-YAA-ARG-|-2AA BONDS.
CC      WHERE XAA CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
CC      - COFACTOR: PROBABLY CALCIUM-DEPENDENT (BY SIMILARITY).
CC      - TISSUE SPECIFICITY: HIGH EXPRESSION IN THE ANTERIOR PITUITARY AND
CC      IN SEVERAL BRAIN REGIONS, THE ATRIUM, AND THE VENTRICLES.
CC      - DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE
CC      ASSISTING THE FOLDING OF THE ZMOGEN WITHIN THE ENDOPLASMIC
CC      RETICULUM.
CC      - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC      - SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.

```